„Cu o simplitate de-a dreptul înviorătoare, Hannah Fry reuşeşte să explice unele dintre cele mai spinoase probleme ale lumii informatice moderne, de la deciziile luate de inteligenta artificială, reţelele neurale şi învăţarea automată, la scandalul stârnit de implicarea firmei britanice Cambridge Analytica în alegerile americane din 2016, neajunsurile automobilelor autonome şi multe altele. Această carte ne demonstrează tuturor de ce este imperios necesar să existe oameni de ştiinţă excelenţi precum Hannah Fry, care să se dedice, în acelaşi timp, şi scrisului”.

KATYGUEST, *Guardian*

HANNAH FRY

HELLO.

WORLD!

CORINT FUTURE

Colecţie coordonată de ADA ROSETI

HANNAH FRY

HELLO.

WORLD!

Revoluţia informatica şi viitorul omenirii

Traducere din limba engleză ADINA IHORA

Cuvânt-înainte CIPRIANI. TOMOIOAGĂ

CORINT BOOKS 2019

Lectură de specialitate: Ciprian I. Tomoioagă

Redactare: Alina Sârbu, Alina Bâltâc

Glosar şi note: Alina Sârbu, Ciprian I. Tomoioagă

Tehnoredactare: Zamfir Adriana Nicoleta PFA

Design copertă: Dan Mihalache

Ilustraţia copertei: e Shutterstock

Ilustraţii interior: p. *27*, *Maşina-câine*, fotografie reprodusă cu permisiunea lui Danilo Vasconcellos Vargas, Universitatea Kyūshū, Fukuoka, Japonia; p. 122, *Gorilă-n piept*, fotografie reprodusă cu permisiunea lui Trafton Drew, Universitatea din Utah, Salt Lake City, SUA; p. 208, *Imaginile lui Steve Talley* e Steve Talley (stânga) şi FBI; p.

211, *Neil Douglas şi sosia lui*, fotografie reprodusă cu permisiunea lui Neil Douglas; p. 216, *Ochelarii „Ţestoasă”*, fotografie reprodusă cu permisiunea lui Mahmood Sharif, Universitatea Carnegie Mellon, Pittsburgh, SUA; p. 216, Georges Biard, *Milla Jovovich la Festivalul de Film de la Cannes*, fotografie reprodusă cu permisiunea autorului.

HANNAH FRY

Hello World: *Being Human in the Age of Algorithms*

Copyright © Hannah Fry Limited 2018 All rights reserved.

Toate drepturile asupra ediţiei în limba română aparţin EDITURII CORINT BOOKS.

ISBN: 978 – 606 – 793 – 598 – 1

Descrierea CIP a Bibliotecii Naţionale a României

FRY, HANNAH

Hello, world!: revoluţia informatică şi viitorul omenirii / Hannah Fry; trad. din lb. engleză: Adina Ihora; cuv. -înainte şi lectură de specialitate: Ciprian I. Tomoioagă.

— Bucureşti: Corint Books, 2019 Index

ISBN 978 – 606 – 793 – 598 – 1

I. Ihora, Adina (trad.)

II. Tomoioagă, Ciprian I. Pref.)

CUVÂNT-ÎNAINTE

Vi s-a întâmplat, uneori, să ajungeţi acasă şi să simţiţi că aveţi chef de stat pe Facebook sau pe Instagram, fără să ştiţi ce anume aţi vrea să vedeţi acolo? Sau să deschideţi o asemenea aplicaţie, doar ca să o închideţi în mai puţin de două minute? Oare cum am dobândit astfel de automatisme?

În capitolul al doilea al acestei minunate introduceri în lumea algoritmilor pentru ne-programatori[1], Hannah Fry

1 Deşi cartea de faţă oferă detalii despre algoritmi, aceasta nu este legată în niciun fel de opera foarte apreciată de programatori Introduction to Algorithms („Algoritmii. 0 introducere”, MIT Press & McGraw Hill, Cambridge, Massachussets, 1990) de Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest şi Clifford Stein, care serveşte drept manual de programare la cea mai prestigioasă universitate politehnică din lume, MIT (Massachussetts Institute of Technology).

oferă mai mult decât răspunsuri la amintita întrebare. Folosind explicaţii simple, însoţite de exemple concise, ea reuşeşte să scoată la iveală cele mai grave probleme pe care internetul şi universul on-line ni le-au creat recent. De la faimoasele scandaluri legate de confidenţialitate ale companiei Facebook şi până la practicile obscure ale magazinelor (on-line şi off-line), aceste istorioare ne ajută să înţelegem dinamicile complexe care sunt folosite azi pentru a ne face să ne cheltuim cea mai preţioasă monedă: atenţia.

Citind sus-amintitul capitol, am trecut printr-un amalgam de stări emoţionale. Întâi, am fost bucuros că autoarea aduce în atenţia publicului larg probleme despre care eu, ca informatician, sunt conştient, dar nu am puterea de a le mediatiza. Apoi, am fost uluit să aflu că unele practici, precum utilizarea cârdului de fidelitate la un magazin, au apărut cu mult înaintea internetului aşa cum îl ştim azi (1993). Am fost entuziasmat şi speriat, în acelaşi timp, când am văzut că algoritmul de recomandare al magazinului Target a depistat sarcina unei adolescente înaintea familiei acesteia, şi că asta se întâmpla în anul 2002! Entuziasmul s-a pierdut şi a rămas doar îngrijorarea când am realizat că, la câte reacţii de *Apreciere*, *Râs*, *Plâns* şi *Furie* am lăsat pe Facebook, platforma de socializare a ajuns să mă cunoască mai bine decât propriii părinţii! Iar îngrijorarea s-a transformat apoi în furie, aflând că analiştii folosesc modelul meu psihologic pentru a-mi influenţa opiniile politice, şi m-a cuprins deznădejdea când am aflat că este imposibil, în momentul de faţă, „să resetezi” ce ştie internetul despre tine, aşa cum a demonstrat-o cazul unei mame care a pierdut sarcina mult dorită, dar nu a reuşit să scape de reclamele destinate gravidelor, care o hăituiau peste tot Online. În final, mi-am recăpătat speranţa că nu e totul pierdut şi încă putem lua măsuri, cum este legea cu privire la protecţia datelor în Europa. Trebuie doar să nu stăm cu mâinile în sân şi să fim conştienţi că datele noastre personale sunt, în ziua de azi,

mai preţioase decât petrolul. Cartea aceasta ne ajută să înţelegem că suntem, fiecare în parte şi toţi împreună, foarte, foarte bogaţi şi puternici. Depinde de noi, doar să le cerem parlamentărilor legi rezonabile, care să ne protejeze, fără să ne îngrădească.

În ciuda sentimentelor diverse prin care ne poate purta, *Helio*, *world*! nu se vrea a fi un articol à la Buzz Feed[2]. Dimpotrivă, am apreciat mult numeroasele anecdote care te ajută să înţelegi şi să te raportezi şi afectiv la chestiunile prezentate, întrucât nu cred că e cineva expert în toate cele şapte domenii abordate. În plus, este evident spiritul de cercetătoare al autoarei din faptul că absolut toate anecdotele sunt cazuri reale, cu conţinut verificabil şi cu surse citate în notele de la final. Ba, uneori, autoarea face trimitere chiar la filmuleţe postate pe YouTube, pe care le-am verificat, ca un tocilar pasionat, şi pot mărturisi că merită efortul de a transcrie linkul în navigator![3] Trebuie observat, însă, că situaţiile prezentate sunt mai mult decât nişte simple poveşti de seară sau subiecte de discuţie la petreceri. Sunt convins că autoarea le-a ales cu grijă pentru a putea aborda răspunsuri diverse la întrebări dificile, pe care, uneori, ne este frică să le punem. Spre exemplu, în dilema vagonului scăpat de sub control[4], nu există un răspuns uşor sau corect; fiecare caz va genera victime. Dar trebuie dat un răspuns, oricât de greu ar fi; iar aceasta se poate face doar înţelegând cum funcţionează maşinile autonome şi ajustându-ne, prin urmare, aşteptările faţă de ele.

Datorită profesiei, sunt eu însumi un creator de algoritmi care, conform intenţiilor mele, vor îmbunătăţi vieţile oamenilor. Motivul pentru care, anterior, menţionam

2 Site de media şi divertisment specializat în articole virale, binecunoscut pentru practicile sale controversate de a atrage cititorii, bazate pe crearea de conţinut emoţionant, dar uneori incorect.

3 Poate mai puţin cel către melodia formaţiei Vanilia, care, amintindu-ne, cumva, de epoca lui „Lasă-mă, papa, la mare!”, este un kitsch melodic ce a fost împins forţat în topurile radio din Marea Britanie.

4 Vezi capitolul „Automobilele autonome”.

uimirea provocată de anumiţi ani este că, în domeniul algoritmilor, lucrurile evoluează într-un ritm ameţitor. Inovaţii care cu doar trei ani în urmă făceau senzaţie pentru că au întrecut performanţa umană, azi sunt considerate învechite şi au fost întrecute de câteva ori. Din acest punct de vedere, după cum recunoaşte şi autoarea, opera de faţă va deveni, poate, incompletă de îndată ce va ieşi din tipografie. Totuşi, consider că *Hello, world*! reprezintă cel mai cuprinzător studiu al impactului pe care îl au algoritmii. După cum veţi vedea, cele şapte capitole tratează influenţa algoritmilor în tot atâtea domenii, care sunt considerate, în principal, tipic umane, spre exemplu medicina, justiţia şi chiar arta. Şi în nenumărate rânduri sunt prezentate cazuri în care algoritmii surclasează oamenii. Dar, oare, înseamnă asta că ne pot înlocui? M-am oprit deseori din citit pentru a reflecta la acest lucru.

Se spune, în ziua de azi, că inteligenţa artificială provoacă deja cea de-a patra revoluţie industrială[5], iar lideri din întreaga lume îşi fac griji în legătură cu impactul pe care îl va avea asupra societăţii, oferind ca motiv principal viteza cu care aceasta se desfăşoară. Uşurinţa cu care companiile vechi şi noi se pot extinde la scară globală le conferă un avantaj fără precedent, iar legislatorii se chinuie să ţină pasul cu tehnologia. În plus, aplicabilitatea largă a acestor algoritmi, care include, după cum am menţionat, domenii tipic umane, inspiră motive de îngrijorare cu privire la locurile noastre de muncă. Am asistat de curând la o prezentare a lui Garry Kasparov care, după cum veţi afla din primul capitol, are motive foarte întemeiate să poarte pică maşinilor cu performanţe super-umane. Am fost surprins să aflu că, în schimb, el este optimist şi susţine ideea dezvoltării unor maşini din ce în ce mai performante. „Când au apărut tractoarele, spunea el, vuiau ziarele de ştiri bucuroase că o astfel de

5 Forumul Economic Mondial, https://www.wefomm.org/agenda/ 2016/01 /the-fourth-industrial-revolution-what-it-means-and-how- to-respond (link accesat cel mai recent la 19 martie 2019).

maşină poate înlocui munca asiduă a 30 de oameni. Nimeni nu era ofensat că un tractor era mai puternic decât el. Ba, mai mult, oamenii se înghesuiau să înveţe să le conducă. De ce lucrurile ar fi diferite acum?”

Este îndreptăţit, oare, Garri să fie atât de optimist în privinţa algoritmilor? Vă invit să vă formaţi propria opinie cu ajutorul acestei lucrări, care va sta, cred eu, la temelia înţelegerii puterii maşinilor din secolul al XXI-lea şi, totodată, a evaluării failibilităţii oamenilor.

CIPRIANI. TOMOIAGĂ

Dedicată Mariei Fry.

Îţi mulţumesc pentru că nu ai acceptat niciodată să ţi se spună „nu”!

NOTA DESPRE TITLU

Când aveam şapte ani, tata ne-a adus, mie şi surorilor mele, un cadou: un ZX Spectrum, un mic computer de numai opt biţi. Pentru prima oară, aveam un calculator numai şi numai al nostru! Probabil că era deja perimat de vreo cinci ani în momentul în care a ajuns la noi, însă, chiar dacă era doar un computer la mâna a doua, am simţit imediat că maşina[6] aceea simpatică era un lucru cu totul şi cu totul minunat. Spectrum era, oarecum, echivalentul unui Commodore 64 (cu toate că numai copiii fiţoşi din cartier aveau aşa ceva), deşi eu am considerat întotdeauna că Spectrum era o drăcovenie mult mai frumoasă. Puteai ţine foarte comod în mâini carcasa netedă din plastic negru, iar tastele din cauciuc gri şi dunga în culorile curcubeului aşezată în diagonală peste unul dintre colţuri îi dădeau un aer destul de prietenos.

În ceea ce mă priveşte, sosirea acelui ZX Spectrum a însemnat începutul unei veri memorabile, pe care am petrecut-o în mansardă, împreună cu sora mea mai mare, programând una pentru cealaltă jocuri de-a

6 Machine (engl.), „(aici) computer”. Aici şi peste tot în continuare am optat să traducem acest termen în conformitate cu definiţia din Traian Anghel, Dicţionar de informatică (Editura Corint, Bucureşti, 2017): „termen echivalent cu calculator, gazdă sau sistem” (pentru o definiţie mai detaliată, vezi glosarul, p. 318) (n. red.).

spânzurătoarea sau desenând forme geometrice simple cu ajutorul unui cod.

Toate aceste chestiuni „avansate” au venit, însă, mai târziu. Pentru început, trebuia să stăpânim bine noţiunile de bază.

Privind în trecut, nu-mi amintesc cu precizie momentul în care am scris primul meu program de computer, dar sunt destul de sigură că ştiu despre ce era vorba. E posibil să fi fost acelaşi program simplu pe care am ajuns să-l predau tuturor studenţilor mei de la University College London (UCL), acelaşi pe care-l veţi găsi pe prima pagină a oricărui manual de introducere în informatică[7]. Pentru că s-a împământenit o tradiţie printre toţi cei care învaţă să programeze, una care e, aproape, un fel de ritual de iniţiere. Prima însărcinare a unui novice în ale computerelor este să programeze calculatorul să afişeze pe ecran o expresie bine-cunoscută:

HELLO, WORLD!

Este o tradiţie care datează încă din anii 1970, când Brian Kemighan a inclus-o ca lucrare practică în manualul său de programare extrem de cunoscut.1 Cartea – de unde provine expresia – a marcat un moment important în istoria calculatoarelor. Microprocesorul tocmai îşi făcuse apariţia în scenă, anunţând transformarea computerelor din ceea ce fuseseră în trecut – nişte maşini grozave pentru specialişti, alimentate cu cartele şi benzi perforate – în ceva mai asemănător cu PC-urile pe care le cunoaştem, cu ecran, tastatură şi cursor care clipeşte. *Hello, world*! şi-a făcut apariţia în momentul în care a devenit o posibilitate să comunici cu computerul.

Ani buni mai târziu, Brian Kemighan i-a relatat unui redactor al revistei *Forbes* ce l-a inspirat să aleagă această expresie. Văzuse un desen animat cu un ou şi cu un puişor abia ieşit din găoace, care, în clipa în care se născuse, ciripise cuvintele: *Hello, world*! şi imaginea acesta îi

7 Pentru definiţiile unor termeni de specialitate, aici şi în continuare, vezi glosarul, p. 315 (n. red.).

rămăsese în minte.

Nu se ştie prea bine cine era puişorul în tot acest scenariu: tinerii care îşi anunţau, triumfători, pătrunderea curajoasă în universul programării computerelor? Sau PC-ul însuşi, trezit din moţăiala sa plictisită printre foi de calcul şi documentetext, pregătit să-şi conecteze mintea la lumea reală şi să îndeplinească cererile noilor săi stăpâni? Poate că amândouă… Cert e, însă, că aceasta rămâne, cu siguranţă, o expresie care îi uneşte pe toţi programatorii şi îi pune în legătură cu fiecare computer care a fost programat vreodată.

Şi trebuie să mărturisesc că îmi mai place ceva la expresia aceasta – ceva ce nu a fost niciodată mai relevant sau mai important decât este în prezent. Căci, în timp ce algoritmii informatici ne controlează şi ne decid viitorul într-o măsură din ce în ce mai mare, *Hello*, *world*! ne reaminteşte de un moment în care între om şi computer se purta un dialog. Ne reaminteşte de clipa în care graniţa care îl desparte pe cel care controlează de cel controlat devine, practic, insesizabilă. Şi marchează, totodată, începutul unui parteneriat – un parcurs comun de explorare a posibilităţilor, în care unul nu poate exista fără de celălalt.

Iar, în era maşinilor, acesta este un sentiment de care merită să ţinem seama.

INTRODUCERE

ORICINE A VIZITAT VREODATĂ parcul naţional Jones Beach din Long Island, statul american New York, probabil că a trecut cu automobilul, în drum spre ocean, pe sub câteva poduri. Acestea au fost construite, iniţial, în folosul celor care intrau sau ieşeau de pe autostradă şi au o caracteristică neobişnuită. Arcuindu-se uşor pe deasupra autostrăzii, ele se află la o înălţime extrem de mică de aceasta, lăsând uneori o distanţă de numai 2, 74 metri până la asfalt.

Designul ciudat are o explicaţie, însă. În anii 1920, Robert Moses, un urbanist bogat şi puternic din New York,

a dorit ca recent terminatul şi foarte premiatul parc-rezervaţie de la Jones Beach să le fie accesibil numai şi numai americanilor albi înstăriţi. Şi, cum ştia prea bine că toţi clienţii săi preferaţi aveau să se deplaseze până la plajă în autoturismele lor personale, în vreme ce locuitorii de culoare din cartierele sărace n-aveau să poată ajunge acolo decât cu autobuzul, el a încercat, în mod intenţionat, să le limiteze celor din urmă posibilitatea de a merge acolo, construind sute de poduri joase de-a lungul autostrăzii. Atât de joase, încât autobuzelor înalte de 3, 5 metri să le fie imposibil să treacă pe sub ele.1

Podurile acestea de-a dreptul rasiste nu sunt singurele obiecte care, deşi neînsufleţite, au reuşit să exercite un control tăcut, clandestin, asupra fiinţelor umane. Istoria este pângărită de nenumărate obiecte şi invenţii care posedau o putere superioară scopului lor declarat.2 Uneori, aceasta fusese integrată, în mod deliberat şi răuvoitor, în însuşi conceptul lor, de la bun început, dar alteori nu a fost decât rezultatul unor omisiuni nechibzuite: gândiţi-vă un pic numai la inexistenţa, în unele zone urbane, a rampelor de acces pentru persoanele ţintuite în scaune cu rotile! Iar câteodată, a fost, pur şi simplu, o consecinţă neintenţionată, aşa cum a fost cazul războaielor de ţesut din secolul al XIX-lea, care fuseseră fabricate pentru a uşura obţinerea unor materiale textile complexe, însă, în cele din urmă, au ajuns să aibă un impact uriaş asupra salariilor, şomajului şi condiţiilor de lucru, încât au devenit, în mod indiscutabil, mai tiranice decât orice capitalist din epoca victoriană.

Invenţiile moderne nu sunt, nici ele, diferite. Întrebaţi-i numai pe locuitorii din orăşelul Scunthorpe[8] din nordul Angliei ce probleme au avut când li s-a interzis să-şi

8 în numele orăşelului se distinge cuvântul cunt (engl.), „(trivial, peiorativ şi ofensator) vagin”. în limbajul curent a fost introdusă expresia Scunthorpe problem, care desemnează o dificultate ce apare atunci când filtrele de internet blochează nume legitime pentru că acestea conţin anumite combinaţii de litere, care, scoase din context, pot fi considerate obscene (n. tr.).

deschidă conturi pe AOL după ce uriaşul furnizor de internet a pus la punct un nou filtru, care excludea vulgarităţile şi care a avut obiecţii faţă de însăşi denumirea localităţii lor.3 Sau pe Chukwuemeka Afigho, nigerianul care a descoperit că un dozator automat pentru săpun lichid elibera soluţia, în mod impecabil, atunci când prietenul său alb îşi punea mâinile sub aparat, însă refuza, cu obstinaţie, să facă acelaşi lucru în cazul său, nerecunoscând nuanţa mai închisă a pielii sale.4 Sau pe Mark Zuckerberg, care, atunci când concepea codul pentru Facebook în camera lui de student de la Harvard, în 2004, nu şi-ar fi imaginat niciodată că propria creaţie avea să fie acuzată că a înlesnit manipularea voturilor în alegeri desfăşurate pe întregul mapamond.5

În spatele fiecăreia dintre aceste invenţii stau algoritmii. Fragmente invizibile de cod care alcătuiesc mecanismele şi rotiţele maşinilor epocii moderne, algoritmii au oferit lumii totul, de la fluxurile de ştiri ale mass-media şi până la motoarele de căutare, de la navigarea prin satelit şi până la programele de recomandare a muzicii, făcând tot atât de mult parte din infrastructura noastră modernă pe cât au făcut, cândva, şi podurile, clădirile sau fabricile. Există algoritmi în spitale, în sălile de judecată şi în autovehiculele pe care le conducem. Tot ei sunt folosiţi de forţele de poliţie, de către lucrătorii din supermarketuri, dar şi de către angajaţii studiourilor cinematografice. Algoritmii au învăţat ce ne place şi ce ne displace; ei ne spun ce să urmărim, ce să citim şi cu cine să avem o întâlnire romantică. Iar, în tot acest timp, au ajuns să deţină puterea ascunsă de a schimba, încet-încet şi în mod subtil, regulile în privinţa a ceea ce înseamnă să fii om.

În această carte vom descoperi marea varietate de algoritmi pe care ne bazăm din ce în ce mai mult, poate chiar fără să ne dăm seama. Ne vom îndrepta atenţia asupra a ceea ce pretind ei că fac, le vom analiza puterea nedeclarată şi vom încerca să ne confruntăm cu întrebările

rămase fără răspuns pe care le ridică. Vom face cunoştinţă cu algoritmii folosiţi de către poliţişti pentru a decide cine să fie arestat, pe aceia care ne fac să alegem dacă să le protejăm pe victimele infracţionalităţii sau să credem în inocenţa celor care sunt puşi sub acuzare. Vom afla care sunt algoritmii pe care îi folosesc judecătorii pentru a stabili sentinţele date infractorilor care au fost găsiţi vinovaţi, adică acelea care ne somează, practic, să decidem cum ar trebui să fie sistemul nostru judiciar. Vom descoperi că există algoritmi pe care medicii îi folosesc pentru a-şi reexamina sau chiar pentru a-şi infirma propriile diagnostice, algoritmi în interiorul autoturismelor fără şofer care tot insistă să ne reevaluăm propria moralitate, dar şi unii care cântăresc modul în care ne exprimăm emoţiile, precum şi alţii care deţin puterea de a ne submina democraţiile.

Nu susţin, nici pe departe, că algoritmii în sine ar fi nocivi. Aşa cum veţi vedea în paginile următoare, există nenumărate motive pentru care merită să rămânem pozitivi şi să privim cu optimism spre ceea ce ne aşteaptă în viitor. Tot ce contează este modul în care sunt utilizaţi algoritmii. GPS-ul a fost inventat pentru lansarea rachetelor nucleare, iar acum ajută la livrarea pizzei, câtă vreme s-a descoperit că repetarea la nesfârşit a unei melodii pop poate fi folosită ca metodă de tortură. Totodată, să ştiţi că, oricât de frumos va fi fost împletită o ghirlandă de flori, dac-aş dori neapărat acest lucru, chiar aş putea folosi una pentru a strangula pe cineva! Să îţi formezi o opinie în privinţa unui algoritm înseamnă să înţelegi relaţia care există între om şi maşină. Fiecare algoritm este, în mod obligatoriu, legat de persoanele care l-au conceput şi îl utilizează.

Ceea ce înseamnă că, în esenţă, aceasta este o carte despre oameni. Despre felul în care suntem, despre direcţia în care ne îndreptăm, despre ceea ce este important pentru noi şi despre felul în care toate acestea sunt transformate prin intermediul tehnologiei. Este o

carte despre relaţia pe care o avem cu algoritmii care există deja, cu aceia care lucrează împreună cu noi, ne potenţează abilităţile, ne corectează greşelile, ne rezolvă problemele, deşi ne mai creează şi altele noi pe parcurs.

Aici ne vom pune întrebări precum: aduc sau nu algoritmii beneficii certe societăţii? Când ar trebui să ai mai multă încredere într-o maşină decât în propria putere de judecată şi când ar trebui să rezişti tentaţiei de a lăsa computerele să comande? Vom descompune algoritmii şi le vom descoperi limitările, iar, în cele din urmă, ne vom cerceta cu atenţie pe noi înşine şi le vom descoperi pe ale noastre. În această carte vom încerca să separăm ce e rău de ce e bine şi să hotărâm în ce fel de lume dorim să trăim.

Pentru că viitorul nu e ceva ce ni se întâmplă, pur şi simplu. Noi suntem aceia care avem puterea de a-l crea.

Capitolul 1
PUTEREA

Garry Kasparov ştia exact cum să-şi intimideze adversarii.

La 34 de ani, era cel mai mare jucător de şah care existase vreodată, având o reputaţie suficient de redutabilă, încât să îl facă pe orice rival al său să fie cuprins de nelinişte. Chiar şi aşa, apela la un şiretlic enervant pe care adversarii săi ajunseseră să-l urască. În timp ce erau aşezaţi, asudând să câştige ceea ce era, probabil, cea mai grea partidă din viaţa lor, rusul îşi lua, nonşalant, ceasul de mână de unde fusese pus, lângă tabla de şah, şi şi-l fixa din nou la încheietură. Era semnalul pe care îl recunoştea toată lumea, acela care însemna că Garry Kasparov se plictisise să se joace cu adversarul său. Ridicarea ceasului semnala că era timpul ca rivalul său să abandoneze partida. Acesta putea să refuze, însă Kasparov avea să-l învingă oricum în curând, iar acest lucru era inevitabil1.

Dar atunci când Deep Blue al lui IBM[9] l-a înfruntat pe

9 în mai 1997, Gări (sau Garry) Kasparov a jucat la New York un meci de şase partide cu super-computerul IBM, Deep Blue. Evenimentul a

Kasparov, în meciul celebru din mai 1997, maşina era imună la asemenea tactici. Rezultatul partidei este bine-cunoscut, însă povestea modului în care Deep Blue şi-a asigurat victoria este mai puţin ştiută. Acea triumfare simbolică a computerului asupra fiinţei umane, care a marcat, în multe privinţe, începutul erei algoritmilor, nu s-a datorat doar simplei puteri de calcul superioare. Pentru a-l învinge pe Kasparov, Deep Blue fusese nevoit să-l înţeleagă nu doar ca pe un procesor extrem de eficient pentru calcularea unor mutări de şah senzaţionale, ci şi ca pe o fiinţă umană.

Pentru început, inginerii de la IBM au luat geniala decizie de a-l proiecta pe Deep Blue astfel încât să pară mult mai nesigur decât era în realitate. Pe durata celebrului meci format din şase partide, maşina s-a abţinut, din când în când, să anunţe ce mutarea dorea să efectueze după ce îşi terminase deja calculele, uneori chiar şi pentru câteva minute bune. Din perspectiva lui Kasparov, întârzierile păreau să lase de înţeles că maşina se căznea, chinuindu-se să efectueze din ce în ce mai multe calcule. Asta părea să confirme ceea ce Kasparov credea că ştia: şi anume, că reuşise să aducă jocul într-o situaţie în care numărul posibilităţilor era atât de năucitor de mare, încât Deep Blue nu putea să ia o hotărâre judicioasă2. În realitate, acesta stătea în repaus, ştiind exact ce piesă să mute şi lăsând, pur şi simplu, timpul să se scurgă. Era un şiretlic plin de viclenie, dar a funcţionat. Chiar din prima partidă a meciului, Kasparov a început să se lase din ce în ce mai distras, tot încercând să aprecieze cât de priceput la şah putea fi Deep Blue3.

Deşi Kasparov a câştigat prima partidă, de la a doua, Deep Blue a reuşit să i se vâre pe sub piele, desluşindu-i modul de a gândi. Kasparov a încercat să-l ademenească într-o capcană, ispitindu-l să captureze câteva piese, în timp ce el însuşi se pregătea – cu câteva mişcări în avans –

fost urmărit de milioane de oameni, în timp real, pe site-ul IBM. Dintre cele şase partide desfăşurate pe parcursul a şapte zile, Kasparov nu a reuşit să învingă calculatorul decât în prima partidă (n. tr.).

să-şi elibereze regina şi să lanseze un atac4. Toţi şahiştii care urmăreau partida, ca şi Kasparov însuşi, se aşteptau ca Deep Blue să muşte momeala şi să cadă în capcană. Numai că, nu se ştie cum, computerul a mirosit el ceva. Spre uimirea lui Kasparov, Deep Blue şi-a dat seama ce punea la cale marele maestru şi i-a blocat acestuia regina, distrugându-i, astfel, orice şansă de a obţine o victorie umană5.

Kasparov a fost, în mod vizibil, oripilat. Faptul că se înşelase în legătură cu ceea ce era computerul capabil să facă l-a terminat, pur şi simplu. Câteva zile mai târziu, într-un interviu luat după meci, el a apreciat că Deep Blue jucase „dintr-odată, pentru o clipă, ca un zeu6”. Mulţi ani mai târziu, reflectând la cum se simţise atunci, Kasparov a scris că „făcuse greşeala să creadă că mutările pe care era surprinzător ca un calculator să le facă reprezentau, totodată, şi nişte mutări grozave, imbatabile, în mod obiectiv”.7 în orice caz, geniul algoritmului triumfase. Felul în care reuşise să înţeleagă mintea omenească şi propensiunea acesteia spre greşeală îl ajutase să îl atace şi să îl învingă pe geniul cel mult prea uman.

Demoralizat, Kasparov a preferat să abandoneze a doua partidă, în loc să lupte pentru a obţine o remiză. Din acel moment, încrederea lui în sine a început să se destrame.

A treia, a patra şi a cincea partidă s-au soldat cu remize. Când s-a ajuns la a şasea, Kasparov era deja terminat. Scorul final al meciului: Deep Blue, *3 ¥2*, Kasparov, 2V2.

A fost o înfrângere ciudată. Kasparov era mai mult decât capabil să iasă din toate poziţiile dificile în care îi ajunseseră piesele de pe tabla de şah, însă el subestimase capacitatea algoritmului, iar apoi se lăsase intimidat de acesta. Reflectând la acel meci, el avea să scrie în 2017:

Am fost atât de impresionat de felul în care juca Deep Blue! Am devenit atât de preocupat de ceea ce era capabil să facă, încât nu am mai izbutit să-mi dau seama că

toate problemele mele se datorau mai mult faptului că eu jucam prost, decât faptului că el juca prea bine.8

Aşa cum vom vedea în repetate rânduri în această carte, aşteptările, speranţele noastre, sunt importante. Povestea înfrângerii unui mare maestru de şah de către Deep Blue arată că puterea unui algoritm nu se limitează neapărat la ceea ce conţin liniile sale de cod. Înţelegerea propriilor defecte şi slăbiciuni – precum şi pe acelea ale computerului – reprezintă cheia pentru deţinerea controlului.

Dar dacă un geniu precum Kasparov nu a înţeles acest lucru, ce speranţe mai avem noi, toţi ceilalţi? în aceste pagini vom vedea cum algoritmii s-au strecurat, practic, în absolut toate aspectele vieţii moderne – de la serviciile medicale şi lupta împotriva infracţionalităţii, până la transport şi politică. Am reuşit, cumva, în decursul timpului, şi să-i ignorăm, şi să ne lăsăm intimidaţi de ei, dar şi să ne minunăm de capacităţile lor, toate în acelaşi timp. Rezultatul este că nu prea avem habar cât de multă putere le conferim şi nici dacă nu cumva am îngăduit deja ca lucrurile să meargă prea departe.

ÎNAPOI LA ELEMENTELE FUNDAMENTALE

Înainte de a ajunge să discutăm despre toate acestea însă, ar fi bine, poate, să ne oprim puţin şi să ne întrebăm ce înseamnă, mai exact, un *algoritm*. Este o noţiune care, deşi frecvent utilizată, nu reuşeşte, cel mai adesea, să transmită prea multe. Asta se întâmplă, în parte, deoarece termenul în sine este destul de neclar. Oficial, el este definit după cum urmează9:

algoritm (s.m.), o procedură pas cu pas de soluţionare a unei probleme sau de îndeplinire a unei sarcini, în principal de către un computer.

Asta e tot! Un algoritm reprezintă, pur şi simplu, o serie de instrucţiuni logice care arată, de la început şi până la sfârşit, cum să se îndeplinească o anumită sarcină. Conform acestei definiţii generale, o reţetă pentru prăjituri poate fi considerată un algoritm. La fel şi o succesiune de

îndrumări pe care i le-ai oferi unui străin care s-a rătăcit. Manualele de la IKEA, videoclipurile despre cum să repari diverse chestii de pe YouTube, până şi cărţile de tip *self-help* – teoretic, orice listă de instrucţiuni autonome cu ajutorul cărora se poate atinge un obiectiv specific şi bine definit ar putea fi considerată un algoritm.

Numai că termenul nu este utilizat chiar în acest sens. De regulă, când spunem *algoritmi* ne referim la ceva special: încă mai reprezintă, în esenţă, nişte liste de instrucţiuni care trebuie urmate pas cu pas, dar aceşti algoritmi sunt, aproape întotdeauna, nişte obiecte matematice. Ei preiau secvenţe de operaţiuni matematice şi (folosind ecuaţii, calcule aritmetice sau algebrice, ori chiar analiza matematică, logica şi probabilităţile) le traduc în cod-program. Algoritmii sunt „alimentaţi" cu informaţii din lumea reală, li se atribuie un obiectiv şi sunt puşi la lucru, să calculeze şi să tot calculeze până îşi ating scopul. Algoritmii sunt aceia care fac din informatică o adevărată ştiinţă, iar, în decursul timpului, au facilitat multe dintre cele mai extraordinare creaţii moderne realizate de diverse maşini.

Există un număr aproape incalculabil de algoritmi diferiţi. Fiecare cu propriile obiective, propriile caracteristici, ingenioasele sale ciudăţenii şi deficienţe, aşa că nu s-a ajuns încă la un consens cu privire la cel mai bun mod de a-i clasifica. Dar, în mod general, poate fi util să ne gândim la sarcinile din lumea reală pe care le îndeplinesc algoritmii, grupându-le în patru mari categorii10:

Stabilirea priorităţilor: conceperea unei liste ordonate

Google Search anticipează ce site cauţi, deoarece ierarhizează, stabileşte prioritatea unui site pe care s-ar putea să doreşti să-l accesezi în funcţie de frecvenţa cu care l-ai utilizat în trecut, comparând rezultatele şi clasificându-le. În mod similar procedează Netflix când îţi sugerează filme pe care s-ar putea să-ţi placă să le

urmăreşti sau Tomâom[10] când îţi alege ruta cea mai rapidă. Toate folosesc un proces matematic pentru a ordona marea varietate de alegeri posibile. Deep Blue a fost, în principal, tot un algoritm de stabilire de priorităţi, unul care trecea în revistă toate mutările posibile de pe tabla de şah şi calcula care dintre ele ar putea oferi şansa cea mai bună de a învinge.

Clasificarea: alegerea unei categorii

Când m-am apropiat de vârsta de 30 de ani, am fost bombardată pe Facebook de reclame la inele de logodnă, împodobite cu diamante. Iar, de îndată ce m-am căsătorit, reclamele pentru testele de sarcină mă urmăreau peste tot pe internet. Pentru aceste mici sâcâieli trebuie să le mulţumesc algoritmilor de clasificare. Aceştia, care sunt atât de îndrăgiţi de cei care lucrează în agenţiile de publicitate, rulează în spatele scenei şi te clasifică drept o persoană interesată de lucrurile respective pe baza caracteristicilor personale. (Şi, cu toate că s-ar putea să aibă dreptate, tot este supărător când reclamele la teste de fertilitate îţi tot răsar pe laptop în mijlocul unei şedinţe!)

Există algoritmi care pot să clasifice şi să îndepărteze automat conţinutul inadecvat de pe YouTube, alţii care pot eticheta pozele de vacanţă în locul tău şi algoritmi care îţi pot scana scrisul de mână şi identifica orice semn de pe o pagină drept o literă din alfabet.

Asocierea: găsirea legăturilor

Asocierea se referă la găsirea şi evidenţierea relaţiilor dintre lucruri. Algoritmii pentru site-urile de anunţuri matrimoniale, precum OKC – upid, se bazează pe asociere, ei căutând legături între membri şi sugerând potriviri în funcţie de rezultatele găsite. Motorul de recomandări al companiei Amazon foloseşte o idee similară, făcând asocieri între ceea ce te interesează şi ceea ce i-a interesat pe foştii clienţi. Aceasta a dus la uimitoarea sugestie de cumpărături care i s-a făcut

10 Companie fondată în 1991, cu sediul central în Amsterdam, Olanda, specializată în sisteme de navigare auto şi hărţi (n. tr.).

utilizatorului *Kerbobotat* de pe site-ul Reddit, după ce cumpărase o bâtă de baseball pe Amazon: „Poate vă interesează să cumpăraţi această cagulă?” 11

Algoritmii trebuie adesea să îndepărteze o parte din informaţii pentru a se concentra asupra a ceea ce este important, să separe, adică, semnalul de zgomot. Uneori fac asta *ad litteram*: astfel, algoritmii de recunoaştere a vorbirii, cum sunt cei care rulează în interiorul Siri, Alexa şi Cortana, trebuie, mai întâi, să-ţi filtreze vocea, pentru a o deosebi de zgomotele din fundal înainte de a începe să lucreze la descifrarea spuselor tale. Uneori o fac în sens figurat: Facebook şi Twitter filtrează, în funcţie de activitatea ta anterioară, postările care au legătură cu ceea ce te interesează, pentru a concepe un flux de ştiri (*feed*) personalizat, care ţi se adresează numai şi numai ţie.

Majoritatea algoritmilor sunt astfel concepuţi, încât să execute o combinaţie a sarcinilor de mai sus. Spre exemplu, Über Pool potriveşte, în mod automat, pasagerii potenţiali cu alţii, care se deplasează în aceeaşi direcţie. Odată ce i se dau punctul de plecare şi destinaţia, aplicaţia trebuie să selecteze posibilele rute care te-ar duce acasă, să te asocieze cu alţi utilizatori care se îndreaptă în aceeaşi direcţie şi să te aloce unui grup pe care l-a ales – toate acestea, în timp ce selectează, cu prioritate, rutele care vor obliga şoferul să execute cele mai puţine schimbări de direcţie, pentru a face cursa cât mai eficientă cu putinţă.12

Prin urmare, asta pot face algoritmii. Acum, să vedem cum reuşesc să ducă la îndeplinire toate acestea. Din nou, deşi posibilităţile sunt, practic, infinite, există un mod de a sintetiza lucrurile. Putem considera că abordările utilizate de algoritmi se încadrează, în linii mari, în două paradigme principale pe care le vom întâlni în această carte.

Algoritmi care se bazează pe reguli

Primul tip de algoritmi se bazează pe reguli.

Instrucţiunile lor sunt scrise de un operator uman, sunt directe şi lipsite de ambiguitate. Ne putem imagina că algoritmii aceştia urmează o schemă logică precum reţeta unei prăjituri. Primul pas:

fa asta. Al doilea pas: dacă faci asta, atunci rezultă cealaltă. Ceea ce nu înseamnă deloc că aceşti algoritmi ar fi simpli – există suficient spaţiu pentru construirea unor programe puternice în cadrul amintitei paradigme.

Algoritmi de învăţare automată

Al doilea tip de algoritmi s-a inspirat din modul în care învaţă fiinţele vii. Pentru a oferi o analogie, să ne gândim cum am putea să-l învăţăm pe un câine să ne dea laba. Nu este nevoie să întocmim o listă exactă de instrucţiuni pe care să i le comunicăm. Dresorul nu trebuie decât să îşi stabilească bine în minte un obiectiv (ce doreşte să facă acel câine), precum şi o modalitate oarecare de a-l răsplăti atunci când duce la îndeplinire în mod corect sarcina respectivă. Totul se rezumă, cât se poate de simplu, la consolidarea comportamentului dezirabil şi la ignorarea celui nedorit, şi, desigur, la exersare, astfel încât câinele să ajungă să-şi dea singur seama ce are de făcut. Când vorbim despre algoritmi, echivalentul celor de mai sus este cunoscut sub denumirea de „algoritm de învăţare automată”, care se înscrie în categoria mai largă a „inteligenţei artificiale”, sau IA. Îi furnizezi maşinii informaţii, un obiectiv şi feed-back atunci când se află pe calea cea bună – şi o laşi să descopere singură modalitatea cea mai bună de a-şi atinge scopul.

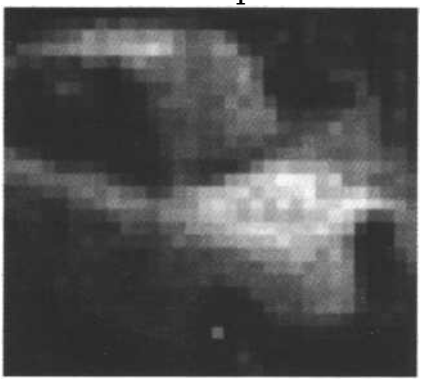
Ambele tipuri de algoritmi au părţi pozitive şi negative. Din moment ce algoritmii care se bazează pe reguli au instrucţiuni scrise de fiinţe umane, ei sunt uşor de înţeles. Teoretic, oricine îi poate descompune şi urmări raţionamentul a ceea ce se petrece înăuntrul lor.13 însă această binecuvântare este, pentru ei, şi un blestem. Algoritmii care se bazează pe reguli vor funcţiona numai pentru rezolvarea problemelor pentru care operatorii umani ştiu cum să scrie instrucţiunile.

Algoritmii de învăţare automată, în schimb, s-au dovedit, recent, a fi remarcabil de buni la rezolvarea unor probleme pentru care scrierea unei liste de instrucţiuni nu este posibilă.

— AI, artificial intelligence (engl.) (n.tr.).

Ei pot recunoaşte obiecte în diverse imagini, ne pot înţelege cuvintele pe măsură ce le rostim şi pot să traducă dintr-o limbă în alta – ceva ce a fost întotdeauna dificil pentru algoritmii care se bazează pe reguli. Dezavantajul este că, dacă vei lăsa o maşină să descopere singură soluţia, procesul pe care îl va urma pentru a-şi îndeplini obiectivul nu va avea sens pentru un observator uman. Căci părţile sale tainice pot fi o enigmă chiar şi pentru cei mai pricepuţi programatori ai zilelor noastre.

Să luăm, de exemplu, sarcina recunoaşterii unei imagini. Un grup de cercetători japonezi a demonstrat de curând cât de bizar i se poate părea unui om modul în care un algoritm „vede” lumea. Poate că aţi mai întâlnit iluzia optică în care nu puteţi spune cu certitudine dacă vă uitaţi la imaginea unei vaze sau la aceea a doua chipuri umane (în caz contrar, aveţi un exemplu în notele de la sfârşitul cărţii14). Iată care ar fi echivalentul acesteia pentru un computer: echipa de cercetători a arătat că simpla modificare a unui singur pixel la roata din faţă a imaginii următoare fusese de ajuns pentru a face ca un algoritm de învăţare automată să se răzgândească şi să considere că fotografia nu mai reprezenta un automobil, ci un câine.15



Pentru unii, ideea ca un algoritm să lucreze fără

instrucţiuni explicite pare să reprezinte cea mai sigură cale spre dezastru. Cum putem controla ceva ce nu înţelegem? Ce se va întâmpla în cazul în care capacităţile maşinilor conştiente de sine şi super-inteligente le vor depăşi pe acelea ale creatorilor lor? Cum am putea garanta că o inteligenţă artificială (IA) pe care nu o înţelegem şi pe care nu o putem controla nu complotează deja împotriva noastră?

Toate acestea sunt întrebări ipotetice interesante şi nu ducem lipsă de cărţi despre ameninţarea iminentă a unei apocalipse cauzate de IA. Îmi cer scuze dacă la aşa ceva vă aşteptaţi, însă cartea de faţă nu este una dintre ele. Cu toate că IA a făcut progrese rapide în ultima vreme, aceasta rămâne „inteligentă" numai în sensul cel mai restrâns al noţiunii. Ar fi, poate, mai util să ne gândim că am asistat la o revoluţie în ceea ce priveşte statistica informatizată, mai degrabă decât la o revoluţie în crearea inteligenţei artificiale. Ştiu că sună mult mai puţin atrăgător (dacă nu cumva *chiar* sunteţi pasionaţi de statistică!), însă este o descriere mult mai corectă a modului în care stau lucrurile în prezent.

Pentru moment, să ne facem griji din cauza unei IA malefice pare, oarecum, ca şi cum ne-am face griji în legătură cu… suprapopularea planetei Marte! * Poate că într-o bună zi vom ajunge şi în punctul în care inteligenţa artificială o va depăşi pe cea umană, însă, pentru moment, nu suntem nici pe departe acolo. La drept vorbind, suntem încă departe de crearea fie şi a unei inteligenţe de nivelul celei pe care o are un arici. Până

— Am parafrazat aici un comentariu făcut de Andrew Ng, specialistul în calculatoare şi pionierul *deep leaming*, într-o conferinţă pe care a susţinut-o în 2015. Vezi Tech Events, GPU Technology Conference 2015, ziua a III-a: „Whafs Next în Deep Leaming", discurs postat pe YouTube (https://www.youtube.com/ watch?v=qP9TOX8T-kI) la 20 noiembrie 2015 (link accesat cel mai recent la 4 martie 2019) (n.a.).

acum, nimeni nu a reuşit să depăşească nivelul inteligenţei unui vierme.[11]

În plus, toată agitaţia stârnită în legătură cu IA reprezintă o abatere a atenţiei de la preocupări mult mai urgente şi – cred eu – de la subiecte mult mai interesante. Haideţi, aşadar, să dăm uitării, pentru moment, omnipotenta inteligenţă artificială şi să ne întoarcem, din viitorul foarte îndepărtat, la timpul prezent, pentru că deja există algoritmi cărora li s-a dat frâu liber să acţioneze şi să ia, în mod autonom, decizii. Unii care stabilesc perioadele de încarcerare ale infractorilor condamnaţi, tratamentele pe care trebuie să le urmeze bolnavii de cancer şi ce trebuie făcut în cazul unui accident rutier. Pentru că algoritmii iau deja, în numele nostru, hotărâri care ne schimbă vieţile.

Întrebarea este: dacă le oferim toată această putere, oare, merită ei toată încrederea noastră?

ÎNCREDERE OARBĂ

Duminică, 22 martie 2009, nu a fost o zi bună pentru Robert Jones. Tocmai îşi vizitase nişte prieteni şi se întorcea cu maşina prin frumosul oraş Todmorden din West Yorkshire, când a observat că i s-a aprins becul indicatorului de benzină al BMW-ului. Mai avea doar 11 kilometri până să găsească o staţie de benzină înainte să i se termine combustibilul, ceea ce era chiar la limită. Din fericire, GPS-ul părea să-i fi găsit o scurtătură şi l-a trimis pe un drumeag îngust şi sinuos care urca pe coasta unui deal.

Robert a urmat indicaţiile maşinii, numai că, pe măsură ce conducea, drumul devenea tot mai abrupt şi mai

11 Simularea creierului unui vierme constituie chiar obiectivul proiectului ştiinţific internaţional OpenWorm. Cercetătorii speră să reproducă, pe cale artificială, reţeaua formată din cei 302 neuroni care se găsesc în creierul unui vierme Caenohabditis elegans. Pentru a pune lucrurile în perspectivă, trebuie menţionat că noi, oamenii, avem în jur de 100 000 000 000 de neuroni. Pentru mai multe informaţii, vezi site-ul proiectului OpenWorm: http://openworm.org/ (link accesat cel mai recent la 4 martie 2019) (n. a.).

îngust, până când, după vreo trei kilometri, s-a transformat într-un adevărat drum de ţară pe care abia-abia dac-ar fi încăput nişte cai, nicidecum automobile! Dar Robert nu a fost impresionat. Conducea câte 8.000 de kilometri pe săptămână ca să-şi câştige pâinea şi ştia cum să se descurce când era la volan. În plus, s-a gândit el, nu avea „niciun motiv să nu aibă încredere în sistemul său de navigare prin satelit Tomâom”. 16

Doar câteva clipe mai târziu, oricine s-ar fi întâmplat să privească în sus din vale ar fi văzut botul BMW-ului lui Robert ivindu-se peste marginea stâncoasă a prăpastiei, fiind salvat de căderea în gol de vreo 30 de metri numai de gardul şubred din lemn de care tocmai se izbise.

A fost nevoie de un tractor şi trei ATV-uri ca să-i recupereze maşina din locul în care o abandonase. Ceva mai târziu în acelaşi an, când s-a prezentat în instanţă sub acuzaţia de conducere imprudentă, Robert a recunoscut că nu se gândise să ignore indicaţiile sistemului său de navigare. „Acesta insista că drumeagul acela era o şosea”, declarase într-un ziar după incident. „Aşa că am avut încredere în el. Nu te aştepţi să fii ghidat drept într-o prăpastie!” 17

Nu, Robert! Nu cred că se aşteaptă cineva la una ca asta!

Istorioara aceasta are, cumva, şi o morală. Cu toate că s-a simţit, probabil, puţin ridicol, pentru că nu a ţinut cont de ceva ce se afla chiar în faţa ochilor săi (cum ar fi, bunăoară, o prăpastie văzută prin parbriz!) şi pentru că i-a atribuit unui algoritm o inteligenţă mai mare decât ar fi meritat, Robert Jones nu era, nici pe departe, singurul om aflat în această situaţie. În fond, Garry Kasparov însuşi căzuse în aceeaşi capcană cu vreo 12 ani în urmă! Şi, în moduri mai puţin spectaculoase, însă la fel de importante, facem aproape toţi această greşeală şi, probabil, nici măcar nu ne dăm seama.

În 2015, oamenii de ştiinţă au început să studieze în ce fel motoarele de căutare precum Google au puterea de

a ne modifica perspectiva asupra lumii.18 Voiau să afle dacă impunem nişte limite sănătoase încrederii pe care o acordăm rezultatelor lor sau dacă nu cumva suntem gata să le urmăm, bucuroşi, chiar dacă ne-ar ghida, metaforic vorbind, peste marginea unei prăpăstii.

Experimentul s-a concentrat asupra alegerilor care urmau să aibă loc în India. Cercetătorii, sub conducerea psihologului Robert Epstein, au selectat 2 150 de alegători indecişi din întreaga ţară şi le-au oferit acces la un motor de căutare special conceput în acest scop, numit „Kadoodle”, pentru a-i ajuta să afle mai multe informaţii despre candidaţi înainte de a hotărî pe care să-l voteze.

Kadoodle era „măsluit”. Fără ca participanţii la studiu să aibă habar, ei fuseseră împărţiţi în grupuri, fiecărui grup fiindu-i arătată o versiune uşor diferită a rezultatelor obţinute de motorul de căutare, care îl favoriza pe unul sau pe altul dintre candidaţi. Atunci când membrii unui grup vizitau site-ul, toate link-urile din partea de sus a paginii favorizau un anumit candidat, ceea ce însemna că ei trebuiau să deruleze pagina în jos, link după link, înainte de a găsi fie şi o singură trimitere către un site care să îi fie favorabilă oricărui alt candidat. Grupuri diferite au fost încurajate să prefere diverşi candidaţi.

Nu este deloc surprinzător că participanţii la experiment şi-au petrecut majoritatea timpului citind site-urile semnalate în partea de sus a primei pagini, pentru că, aşa cum spune gluma aceea răsuflată, cel mai bun loc unde să ascunzi un cadavru este pe pagina a doua a rezultatelor unei căutări pe Google. Aproape niciunul dintre participanţi nu a dat vreo atenţie link-urilor care apăreau mult în josul listei. Cu toate acestea, măsura în care ordonarea influenţase opiniile voluntarilor l-a şocat până şi pe Epstein. După numai câteva minute de privit rezultatele măsluite ale motorului de căutare, când erau întrebaţi cu cine aveau să voteze, participanţii erau mai dispuşi să aleagă candidatul pe care-l favorizase Kadoodle. Cu

uluitorul procent de 12%!

În interviul dat revistei *Science* în 201519, Epstein a explicat ce se întâmpla:

Ne aşteptăm ca motorul de căutare să facă alegeri înţelepte. Participanţii la experiment îşi spun: „Ei bine, da, observ avantajarea unui anumit candidat, iar asta îmi transmite […] că, de fapt, motorul de căutare îşi face datoria.

Dar, dacă luăm în considerare cât de multe dintre informaţiile noastre le primim de la algoritmi precum motoarele de căutare, este încă şi mai ameninţător să aflăm în ce măsură voluntarii credeau că îşi influenţaseră propriile opinii: „Când oamenii nu sunt conştienţi că sunt manipulaţi, ei tind să creadă că au adoptat o părere nouă din proprie voinţă”, scria Epstein în lucrarea sa originală.20

Fireşte, Kadoodle nu este singurul algoritm acuzat că ar manipula, în mod subtil, opiniile politice ale cetăţenilor. Vom insista mai mult asupra acestui fapt în capitolul „Datele personale informatizate”, însă, pentru moment, merită să reţinem ce sugerează experimentul acesta că simţim faţă de algoritmii care au dreptate în majoritatea timpului: sfârşim prin a crede că ei dau întotdeauna dovadă de un raţionament superior.21 în scurt timp, nu mai suntem nici măcar conştienţi de propriul subiectivism faţă de aceştia.

Pretutindeni în jurul nostru, algoritmii oferă un fel de sursă convenabilă de autoritate. O cale uşoară de a delega responsabilitatea; o scurtătură pe care o întrebuinţăm fără să ne mai gândim. Cine, oare, va da clic de fiecare dată pe a doua pagină de rezultate de pe Google şi cine, oare, va analiza critic fiecare rezultat? Sau cine va trece pe la fiecare companie aeriană să verifice dacă Skyscanner chiar afişează zborurile cele mai ieftine? Sau cine va scoate o riglă şi o hartă rutieră ca să aibă confirmarea că GPS-ul chiar îi sugerează drumul cel mai scurt? Eu, una, nu, de asta puteţi fi siguri!

Aici trebuie, însă, să facem o distincţie. A avea încredere într-un algoritm fiabil de obicei este una, iar a avea încredere oarbă într-unul căruia nu-i înţelegi prea limpede caracteristicile este cu totul şi cu totul altceva.

CÂND INTELIGENŢA ARTIFICIALĂ ÎNTÂLNEŞTE PROSTIA VERITABILĂ

În anul 2012, mai multe persoane cu dizabilităţi din Idaho au fost informate că le fusese redusă suma pe care o primiseră până atunci din partea programului federal de asistenţă medicală Medicală.22 Deşi tuturor le fusese recunoscut dreptul de a beneficia de acest ajutor, statul le redusese susţinerea financiară – fără nicio avertizare – chiar şi cu 30%23 lăsându-i pe bieţii oameni să se chinuiască să-şi achite costurile îngrijirilor medicale. Nu era vorba însă de o decizie politică, doar de consecinţa aplicării unui nou „instrument bugetar”, recent adoptat de Idaho Departament of Health and Welfare, ministerul sănătăţii din statul Idaho: un program de software care calcula automat cuantumul ajutorului financiar pe care trebuia să-l primească fiecare bolnav.24

Problema era că deciziile „instrumentului bugetar” nu păreau să aibă prea multă logică. Din câte îşi putea da seama cineva din afara sistemului, cifrele pe care le propunea programul erau, în mare parte, absolut arbitrare. Unor persoane li se dădeau bani mai mulţi decât în anii precedenţi, pe când altele şi-au văzut bugetul redus cu zeci de mii de dolari, astfel încât riscau să fie nevoite să-şi părăsească locuinţele pentru a putea fi îngrijite într-o instituţie publică.25

Incapabili să înţeleagă de ce le fuseseră reduse indemnizaţiile sau să protesteze efectiv împotriva acestor reduceri, oamenii au apelat la American Civil Liberties Union (ACLU) [12] pentru a fi ajutaţi. Cazul lor a fost preluat de Richard Eppink, directorul juridic al filialei ACLU din

12 Uniunea Americană pentru Libertăţi Civile, ONG care luptă, conform propriei carte: „pentru apărarea şi menţinerea drepturilor individuale şi a libertăţilor garantate tuturor cetăţenilor acestei ţări de către Constituţia şi legile Statelor Unite” (n. tr.).

Idaho26, care a ţinut să spună următoarele într-o postare pe blog din 2017:

Am crezut, iniţial, că acest caz avea să fie o chestie simplă şi că tot ce urma să fim nevoiţi să facem era să-i întrebăm pe reprezentanţii statului: „În regulă, ce-ar fi să ne spuneţi şi nouă de ce au scăzut atât de mult sumele astea în dolari?” 27

În realitate, a fost nevoie de patru ani, 4.000 de reclamanţi şi un proces colectiv pentru a se lămuri ce se întâmplase, de fapt.28

Eppink şi echipa sa au început prin a cere amănunte cu privire la modul de funcţionare a algoritmului, dar angajaţii Medicală au refuzat să le explice cum calculaseră sumele respective, argumentând că software-ul care evalua cazurile constituia un „secret comercial”, motiv pentru care nu putea fi făcut public.29 Din fericire, judecătorul care prezida şedinţa nu a fost de acord cu acest lucru. Aşa se face că instrumentul bugetar care exercita atâta putere asupra locuitorilor cu dizabilităţi a fost predat şi s-a dovedit că nu era, nici pe departe, cine ştia ce inteligenţă artificială sofisticată sau model matematic minunat construit, ci un simplu tabel în Excel! 30

În acesta, calculele ar fi trebuit să se bazeze pe istoricul unor cazuri, numai că informaţiile erau atât de înţesate de bug-uri şi erori, încât, în cea mai mare parte, erau absolut nefolositoare.31 Lucru încă şi mai grav, de îndată ce echipa de la

ACLU a reuşit să descâlcească ecuaţiile, au fost descoperite „greşeli statistice fundamentale în modul în care era structurată însăşi formula de calcul”. Instrumentul bugetar produsese efectiv rezultate aleatorii pentru un număr imens de persoane! Algoritmul – dacă îl putem numi astfel – era de o calitate atât de proastă, încât, în cele din urmă, instanţa l-a declarat neconstituţional.32

În acest caz s-au manifestat două şiruri paralele de erori umane. Mai întâi, cineva a scris un tabel în Excel

numai bun de aruncat la gunoi; apoi, alte persoane au avut, cu naivitate, încredere în acesta! De fapt, „algoritmul” nu era decât munca de mântuială a cuiva, ambalată frumos sub forma unui cod. De ce erau, atunci, atât de dornici funcţionarii care lucrau pentru stat să ia apărarea unui lucru atât de tâmpit?

Iată opinia lui Eppink în această chestiune:

Ne-am confruntat cu acea slăbiciune pe care o avem cu toţii când vine vorba de rezultatele obţinute de un computer, pe care, pur şi simplu, nu le mai punem sub semnul întrebării... Când un calculator generează ceva – care se bazează pe munca unui statistician care a analizat anumite informaţii şi a inventat o formulă –, tindem să avem, orbeşte, încredere în formula respectivă, fără să ne mai întrebăm: „Hei, ia stai puţin! Cum funcţionează, de fapt, chestia asta?” 33

Ştiu că să descompui formule matematice pentru a analiza cum funcţionează nu este tocmai distracţia preferată a tuturor oamenilor (chiar dacă este a mea). Cu toate acestea, Eppink a scos în evidenţă o problemă incredibil de importantă cu privire la tendinţa noastră de a da crezare algoritmilor, de a-i lua de buni, fără să ne mai întrebăm cum funcţionează, ce se petrece în spatele faţadei lor.

În anii în care, ca matematician, am lucrat cu informaţii şi algoritmi, am ajuns să cred că singura modalitate de a aprecia cu obiectivitate dacă un algoritm este demn de încredere este să mă lămuresc cum funcţionează. Experienţa pe care o deţin m-a învăţat că algoritmii seamănă destul de mult cu scamatoriile unui magician. La început, par să fie, nici mai mult, nici mai puţin, decât magie pură, dar, de îndată ce ai aflat în ce constă scamatoria, misterul dispare. Adesea, în spatele faţadei se ascunde ceva amuzant de simplu (sau îngrijorător de neglijent, de nesăbuit). Aşa că, în capitolele următoare şi în legătură cu algoritmii pe care îi vom analiza, voi încerca, ori de câte ori îmi va fi cu putinţă, să

vă fac să intuiţi ce se petrece în spatele faţadei acestora. Atât cât să vedeţi cum se fac scamatoriile informatice – chiar dacă nu îndeajuns de amănunţit, încât să le puteţi realiza apoi voi înşivă.

Însă, chiar şi pentru cei mai înrăiţi pasionaţi de matematică, tot vor mai exista situaţii în care algoritmii le vor cere să se arunce, orbeşte, cu capul înainte şi să-i creadă „pe cuvânt”. Poate pentru că, aşa cum este cazul cu Skyscanner sau cu rezultatele căutării pe Google, verificarea modului în care funcţionează respectivele programe nu este posibilă. Sau poate pentru că, aşa cum s-a întâmplat în cazul instrumentului bugetar din Idaho şi în alte situaţii care vor fi prezentate în continuare, algoritmul respectiv este considerat un „secret comercial”. Sau poate pentru că, aşa cum se întâmplă cu anumite procedee de învăţare automată, unei fiinţe umane îi este, pur şi simplu, cu neputinţă să urmărească procesul logic care se desfăşoară în interiorul algoritmului.

Aşadar, vor exista momente în care vom fi siliţi să cedăm controlul necunoscutului, chiar dacă vom fi pe deplin conştienţi că algoritmul s-ar putea să greşească. Momente în care vom fi nevoiţi să alegem între propriul raţionament şi cel al maşinii. Şi când, dacă ne vom hotărî să avem, mai degrabă, încredere în instinctele noastre, nu în calculele sale, va trebui să dăm dovadă de foarte mult curaj.

CÂND SĂ NESOCOTIM ALGORITMII

Stanislav Petrov era un ofiţer din armata rusă responsabil cu supravegherea sistemelor de avertizare nucleară timpurie care protejau spaţiul aerian sovietic. Sarcina lui era să-şi alerteze imediat superiorii în cazul în care computerul indica vreun semn că un atac american era iminent.34

La 26 septembrie 1983, Petrov era la datorie atunci când, după miezul nopţii, sirenele au început să urle. Era semnalul de alarmă de care se temea toată lumea: sateliţii sovietici identificaseră o rachetă inamică, una care se

îndrepta către teritoriul Rusiei! Cum Războiul Rece era în toi, o lovitură armată era cât se poate de plauzibilă, numai că, totuşi, ceva l-a cam pus pe gânduri pe Petrov şi l-a făcut să nu ia nicio măsură. El nu era convins că avea deplină încredere în algoritm, căci acesta identificase numai cinci rachete, ceea ar fi constituit o lovitură de deschidere absurd de redusă pentru un atac nuclear american.35

Petrov a înlemnit în scaunul său. Din clipa aceea totul depindea de el: putea fie să tragă semnalul de alarmă şi să arunce întreaga lume într-un război nuclear aproape sigur, fie să aştepte, ignorând uzanţele şi ştiind că fiecare secundă care se scurgea însemna că, de fapt, conducătorilor ţării sale le rămânea tot mai puţin timp la dispoziţie pentru a lansa o contralovitură.

Din fericire pentru noi toţi, Petrov a hotărât să aleagă cea de-a doua variantă. Nu putea fi absolut sigur că alarma se declanşase din greşeală, dar, după 23 de minute – care trebuie să i se fi părut, atunci, o eternitate –, când a fost limpede că nicio rachetă nucleară nu căzuse pe pământ rusesc, a ştiut că avusese dreptate. Algoritmul fusese acela care greşise.

Dacă sistemul ar fi acţionat complet independent, fără să aibă nevoie de un om ca Petrov care să acţioneze ca arbitru final, istoria s-ar fi desfăşurat, fără îndoială, într-un mod diferit. Rusia ar fi lansat, aproape sigur, ceea ce ar fi considerat a fi un contraatac şi ar fi declanşat, totodată, un război nuclear în toată regula. Dacă putem să tragem o învăţătură din toată povestea, atunci aceasta este că elementul uman trebuie să fie o parte indispensabilă a procesului: să dispui de o persoană cu drept de veto care să fie în măsură să verifice sugestiile făcute de un algoritm înainte de luarea unei decizii este singurul mod raţional de a evita greşelile.

În definitiv, numai oamenii vor simţi că poartă povara răspunderii pentru deciziile pe care le iau. Un algoritm însărcinat cu transmiterea mesajelor până la nivelul

oficialilor de la Kremlin nu ar fi aşteptat nicio secundă să îşi ducă la îndeplinire însărcinarea, pentru că nu ar fi şovăit tot analizând consecinţele posibile ale unei astfel de decizii. Pe de altă parte, ce a făcut Petrov? „Ştiam cât se putea de bine că rumeni nu ar fi reuşit să-mi corecteze greşeala, dacă aş fi făcut-o”. 36

Singura problemă în legătură cu această concluzie este că nici pe fiinţele umane nu te poţi baza chiar întotdeauna. Uneori, asemenea lui Petrov, vor avea dreptate să nu ţină cont de sugestiile unui algoritm. Însă, adeseori, cel mai bine ar fi să ne ignorăm instinctele.

Ca să ofer un alt exemplu din domeniul securităţii, unde situaţiile în care fiinţele umane ignoră în mod greşit un algoritm sunt, din fericire, rare, exact asta s-a întâmplat în cazul celebrei prăbuşiri a montagnes russes-ului denumit The Smiler (Zâmbăreţul”), din Alton Towers, cel mai mare parc de distracţii tematic din Marea Britanie.37

În iunie 2015, doi ingineri au fost chemaţi să remedieze o defecţiune la un montagnes russes. După ce au rezolvat problema, au trimis un vagon gol pe circuit, să verifice dacă totul funcţiona aşa cum trebuie, însă nu au observat că acesta nu s-a mai întors. Din cine ştie ce motiv, vagonul gol s-a răsturnat pe o pantă şi s-a oprit la mijlocul circuitului.

Între timp, fără ştirea inginerilor, personalul de deservire a adăugat la tren un vagon suplimentar pentru a face faţă cozilor din ce în ce mai mari. Odată ce au primit undă verde de la cabina de control, angajaţii au început să umple vagoanele cu pasageri veseli, cărora le-au legat centurile, şi au trimis primul tren pe traseu, fără să aibă habar de existenţa vagonului gol, înţepenit în mijlocul şinelor, care fusese trimis de ingineri şi care stătea de acum chiar în drum.

Din fericire, arhitecţii montagne russes-ului prevăzuseră o situaţie de acest gen, iar algoritmii de siguranţă au funcţionat exact aşa cum fuseseră proiectaţi s-o facă. Pentru a evita o coliziune sigură, trenul ticsit de

oameni a fost oprit automat în vârful primului urcuş, iar un semnal de alarmă s-a declanşat în cabina de control. Numai că inginerii – încrezători că abia reparaseră defecţiunea – au tras concluzia că sistemul automat de avertizare greşea.

Ignorarea algoritmului nu era o acţiune simplă: era necesar ca amândoi să fie de acord şi să apese simultan câte un buton pentru a repomi montagne russes-ul. Procedând în felul acesta, trenul ticsit de oameni a fost trimis la vale şi s-a lovit în plin de vagonul rămas pe şine. Consecinţele au fost de-a dreptul groaznice. În urma impactului, mai multe persoane au suferit leziuni devastatoare şi două adolescente şi-au pierdut picioarele.

Aceste două situaţii de viaţă şi de moarte, semnalele de alarmă declanşate în cazul lui Petrov şi al montagnes russes-ului din parcul de distracţii Alton Towers, ilustrează, în mod tragic, o dilemă profundă: în echilibrul de forţe dintre oameni şi algoritmi, cine (sau ce) ar trebui să aibă ultimul cuvânt?

LUPTA PENTRU PUTERE

Această dezbatere durează de multă vreme. În 1954, Paul Meehl, profesor de psihologie clinică la Universitatea din Minnesota, a supărat o întreagă generaţie… umană atunci când a publicat *Clinical versus Statistical Prediction*[13], o carte în care insista ferm asupra unuia din aspectele argumentaţiei.38

În cartea sa, Meehl a comparat sistematic, riguros, performanţele oamenilor şi ale algoritmilor într-o mare diversitate de situaţii (de la prezicerea notelor studenţilor şi până la evaluarea sănătăţii mintale a pacienţilor) şi a ajuns la concluzia că algoritmii matematici, oricât ar fi de simpli, vor face, aproape întotdeauna, previziuni mai bune decât oamenii.

Nenumărate alte studii, efectuate în jumătatea de secol care a urmat, au confirmat constatările lui Meehl. Dacă sarcina pe care o ai de îndeplinit presupune

13 „Decizia clinică împotriva celei statistice” (n. tr.).

efectuarea oricărui tip de calcul, poţi să pariezi de fiecare dată că algoritmii se vor descurca mai bine: de la stabilirea diagnosticelor medicale sau a vânzărilor prognozate, la prezicerea tentativelor de suicid sau a gradului de satisfacţie în carieră, precum şi atunci când trebuie evaluat absolut orice, de la condiţia fizică necesară pentru îndeplinirea stagiului militar şi până la preconizarea performanţelor ca student.39 Maşina poate că nu va acţiona perfect, dar a-i lăsa omului drept de veto asupra algoritmilor nu va duce, în anumite cazuri, decât la înmulţirea erorilor.[14]

Poate că asta nu ar trebui să ne surprindă. Noi nu am fost concepuţi pentru a face calcule. În fond, când mergem la supermarket, nu dăm peste un şir de casieri care ne verifică din priviri cumpărăturile făcute, pentru a estima cam cât ar costa. Pentru asta avem un algoritm (incredibil de simplu)

care socoteşte în locul nostru. Şi, în majoritatea timpului, ar fi mai bine să lăsăm maşina să calculeze. Este exact ca în gluma aceea care circulă în lumea piloţilor de avion, potrivit căreia, de fapt, cel mai bun echipaj de zbor este compus din trei membri: un pilot, un computer şi un câine. Calculatorul face avionul să zboare, iar pilotul hrăneşte câinele. Patrupedul se află acolo ca să-l muşte pe om dacă acesta ar încerca să se atingă de computer.

Numai că în relaţia noastră cu maşinile se manifestă un paradox. Deşi tindem să acordăm o încredere neţărmurită tuturor lucrurilor pe care nu le înţelegem, de îndată ce ştim că un algoritm poate să facă greşeli, avem

---

14 în mod surprinzător, una dintre rarele excepţii în ceea ce priveşte superioritatea algoritmilor în a face previziuni provine dintr-o serie de studii realizate la sfârşitul anilor 1950 şi pe parcursul anilor 1960, despre „diagnosticarea” (termenul le aparţine cercetătorilor vremii, nicidecum mie!) homosexualităţii. în acele exemple, raţionamentul uman a făcut predicţii mult mai bune, depăşind ca performanţă tot ce reuşea să prevadă algoritmul - ceea ce sugerează că anumite lucruri sunt atât de specifice fiinţei umane, încât informaţiile şi formulele matematice vor avea întotdeauna dificultăţi în a le descrie (n. a.).

obiceiul destul de enervant de a reacţiona exagerat şi de a-l înlătura cu totul, revenind la propriul nostru raţionament deficitar. Cercetătorii cunosc acest comportament, pe care l-au denumit „aversiune faţă de algoritmi”. Oamenii au mai puţină toleranţă faţă de erorile unui algoritm decât faţă de propriile greşeli, chiar dacă acestea din urmă sunt mai mari, mai grave.

Fenomenul a fost demonstrat în repetate rânduri în cadrul unor experimente40 şi, într-o oarecare măsură, îl puteţi recunoaşte şi în propriile dumneavoastră manifestări. Spre exemplu, ori de câte ori Citymapper mă anunţă că, de fapt, călătoria mea va dura mai mult decât m-am aşteptat, nu mă pot abţine să cred că tot eu ştiu mai bine (chiar dacă, de cele mai multe ori, asta înseamnă că voi sfârşi prin a întârzia). Cu toţii am făcut-o pe Siri „idioată”[15] măcar o dată, uitând cumva, pe parcurs, cât de uimitoare sunt realizările tehnologice care au dus la alcătuirea unui asistent personal vorbitor pe care să-l poţi ţine în mână! în perioada de început a utilizării aplicaţiei cu GPS Waze pe mobil, m-am trezit, la un moment dat, prinsă într-un ambuteiaj după ce mă convinsesem singură că urma s-ajung mai repede la destinaţie dacă aveam să o iau pe nişte străduţe lăturalnice, nu pe ruta pe care mi-o indicase Waze. (De-atunci am aflat că niciodată nu aveam dreptate când presupuneam asta.) Acum am ajuns să am încredere în Waze şi – asemenea lui Robert Jones şi a BMW-ului său – o voi urma orbeşte, oriunde m-ar conduce (deşi cred, totuşi, că aş refuza să mă arunc într-o prăpastie).

Tendinţa aceasta de a vedea lucrurile în alb şi negru – de a privi algoritmii fie ca nişte stăpâni omnipotenţi, fie ca pe o grămadă inutilă de gunoaie – este, întrucâtva, o problemă în era noastră high-tech. Dacă ne propunem să profităm la maximum de tehnologie, trebuie să găsim o

15 Deoarece majoritatea utilizatorilor români folosesc varianta americană a acestei aplicaţii, a cărei voce este aceea a actriţei Susan Bennett, am optat ca aici şi în continuare să o considerăm pe Siri de sex feminin (n. red.).

cale de a fi obiectivi. Şi ar fi bine să învăţăm din greşeala lui Kasparov şi să ne recunoaştem greşelile, să punem sub semnul întrebării reacţiile instinctive şi să fim ceva mai conştienţi de sentimentele noastre faţă de algoritmii care au ajuns să ne înconjoare. Pe de altă parte, ar trebui să coborâm algoritmii de pe piedestal, să-i studiem cu ceva mai multă atenţie şi să ne întrebăm dacă sunt, cu adevărat, în stare de tot ceea ce pretind c-ar putea face. Este singura modalitate de a decide dacă merită să deţină toată puterea care le-a fost conferită.

Din nefericire, toate acestea sunt, de multe ori, mai uşor de zis decât de făcut. Deseori nu prea vom avea vreun cuvânt de spus în legătură cu puterea şi răspândirea algoritmilor, chiar şi atunci când vine vorba despre cei care ne afectează în mod direct.

Asta este, mai cu seamă, adevărat pentru algoritmii care fac schimb cu cel mai modern produs de bază: informaţia. Algoritmii care ne urmăresc în tăcere pe internet, cei care culeg informaţiile noastre personale, ne invadează intimitatea şi ne lasă impresia că am avea mână liberă, numai pentru a ne influenţa apoi, în mod subtil, comportamentul. Iar în această adevărată „furtună perfectă" a încrederii, puterii şi influenţei prost plasate, consecinţele au potenţialul de a modifica în mod fundamental societatea în care trăim.

Capitolul 2

DATELE PERSONALE INFORMATIZATE

În 2004, la puţin timp după ce crease Facebook, studentul Mark Zuckerberg a avut următorul schimb de mesaje instantanee, pe chat, cu un prieten:

ZUCK: Mda, aşa că, dac-ai vreodată nevoie de informaţii despre oricine de la Harvard…

ZUCK: Nu trebuie decât să-mi ceri!

ZUCK: Am peste 4.000 de e-mail-uri, poze, adrese…

[NUMELE PRIETENULUI ESTE CENZURAT]: Ce? Cum ai reuşit una ca asta?

ZUCK: Oamenii mi le-au comunicat de bunăvoie.

ZUCK: Habar n-am de ce.

ZUCK: Cic-ar avea „încredere în mine*!

ZUCK: Tâmpiţi ordinari! 1

Ca urmare a scandalului în care a fost implicat Facebook în 2018, aceste cuvinte au fost publicate în mod repetat de jurnalişti, din dorinţa de a insinua că ar fi existat o atitudine machiavelică faţă de confidenţialitatea informaţiilor în cadrul companiei. Eu, una, cred c-am putea da dovadă de mai multă mărinimie atunci când vine vorba să interpretăm comentariile pline de lăudăroşenie ale unui tânăr de 19 ani. Dar mai cred şi că Zuckerberg se înşela. Oamenii nu îi oferiseră, pur şi simplu, informaţii despre ei înşişi. Ci i le puseseră la dispoziţie ca parte a unui schimb. Pentru ele, li se oferea acces la un algoritm care le permitea să ţină legătura, fără restricţii, cu prietenii şi familia. În plus, li se punea la dispoziţie un loc în care să-şi împărtăşească experienţele de viaţă cu alte persoane. Era reţeaua lor privată în imensitatea World Wide Web-ului. Nu ştiu ce părere aveţi, dar, pe vremea aceea, eu, una, în mod cert consideram că era un schimb echitabil.

Există, însă, o singură fisură în raţionamentul acesta: nu suntem întotdeauna conştienţi de implicaţiile pe termen mai lung ale unui asemenea schimb. Ne este, arareori, limpede ce pot face informaţiile noastre sau cât de valoroase pot deveni ele atunci când sunt introduse într-un algoritm inteligent. Şi, în consecinţă, nu ne este clar nici cât de ieftin am fost cumpăraţi.

PUŢIN CÂTE PUŢIN…

Companiile care gestionează supermarketuri au fost printre primele care au recunoscut valoarea strângerii informaţiilor despre o persoană. Într-un domeniu în care firmele concurează permanent pentru a atrage atenţia cumpărătorilor (cu scopul de a adăuga mici preferinţe în comportamentul lor la shopping* care le vor întări, în timp, loialitatea faţă de marca produsului, faţă de brand), fiecare uşoară îmbunătăţire contează, căci, treptat, se adaugă altora similare şi se poate transforma într-un avantaj uriaş.

Acesta a fost motivul care a dus

— *Buying behaviour* (engl.), „comportament de/la cumpărături/ shopping”. Am optat să păstrăm şi anglicismul shopping în traducerea acestei sintagme, deoarece, conform *Dicţionarului explicativ al limbii române* (DEX, ediţia din 2009), el este cuprinzător, descriind mai multe lucruri, nu doar simpla achiziţionare a unor produse, căci este definit drept: „vizitarea magazinelor pentru a vedea vitrinele, galantarele şi a face cumpărături” (n. red.).

la organizarea experimentului revoluţionar din 1993 realizat de supermarketul britanic Tesco.

Sub îndrumarea echipei formate din soţii Edwina Dunn şi Clive Humby şi lansat în magazine anume selectate, Tesco a emis noul său Clubcard – un cârd plastifiat de dimensiunea şi forma unui card de credit, pe care clienţii îl puteau prezenta la casă atunci când îşi achitau cumpărăturile. Schimbul era simplu: la fiecare tranzacţie în care folosea Clubcard, clientul acumula puncte pe care le putea utiliza pentru viitoare cumpărături din magazin, câtă vreme Tesco avea se înregistreze acea vânzarea şi s-o asocieze cu numele clientului.2

Informaţiile culese în acel prim experiment cu ajutorul lui Clubcard au fost extrem de limitate. Pe lângă numele şi adresa clientului, programul nu înregistra decât ce sumă cheltuise acesta şi când anume, nu şi ce produse avusese în coşul de cumpărături. Totuşi, din relativ modesta recoltă de informaţii, Dunn şi Humby au reuşit să extragă nişte concluzii extrem de valoroase.

Ei au descoperit că un număr mic de clienţi fideli realiza un volum mare din vânzări. Au aflat, examinând cod poştal după cod poştal, ce distanţă erau oamenii dispuşi să parcurgă până la magazine. Au văzut care erau cartierele în care concurenţa avea mai mult succes şi cartierele în care avea preponderenţă Tesco. Informaţiile le dezvăluiau ce clienţi reveneau în fiecare zi şi care îşi amânau cumpărăturile pentru sfârşitul săptămânii, înarmaţi cu

aceste cunoştinţe, au trecut la treabă: au încercat să transforme comportamentul la shopping al clienţilor care utilizau Clubcard, trimiţându-le, prin poştă, diverse cupoane. Clienţilor care cheltuiau mai mult li s-au oferit vouchere cu sume cuprinse între 3 şi 30 de lire sterline. Celor care cheltuiau sume mici, li s-a trimis un stimulent mai modest – între 1 şi 10 lire sterline. Iar rezultatele au fost uimitoare. Aproape 70% dintre cupoane au fost răscumpărate, iar în magazine clienţii şi-au ticsit coşurile: persoanele care deţineau un Clubcard au cheltuit, per ansamblu, cu 4% mai mult decât cele care nu deţineau unul.

Pe 22 noiembrie 1994, Clive Humby a prezentat concluziile experimentului său membrilor consiliului de administraţie al companiei Tesco. Le-a arătat toate informaţiile culese, ratele de răspuns, dovezile satisfacţiei clienţilor, creşterile explozive ale vânzărilor. Membrii consiliului au ascultat în tăcere. La sfârşitul prezentării, preşedintele de atunci cil firmei a vorbit primul:

Ceea ce mă sperie în toată chestiunea asta, a spus el, este că voi aţi aflat mai multe informaţii despre clienţii mei în trei luni, decât am aflat eu în 30 de ani! 3

Clubcard-ul a fost extins la toţi clienţii Tesco şi i se recunoaşte, pe scară largă, meritul de a fi ajutat compania să o ia înaintea principalului său rival, Sainsbury’s, şi să devină cel mai mare lanţ de supermarketuri din Marea Britanie. Cu timpul, informaţiile colectate au devenit mai detaliate, ceea ce a făcut ca obiceiurile de shopping ale clienţilor să devină tot mai uşor de urmărit.

Chiar de la începuturile cumpărăturilor on-line, echipa* a introdus elementul cunoscut drept Afy *Favourites*”, în care orice obiect cumpărat cu ajutorul cârdului de fidelitate apărea evidenţiat atunci când clientul se autentifica pe site-ul Tesco. Asemenea Clubcard-ului, şi acest element a avut un succes răsunător. Oamenii puteau găsi cu uşurinţă produsele pe care le doreau, fără să mai navigheze printr-o mulţime de pagini. Vânzările au crescut,

iar clienţii au fost mulţumiţi.

— Dunnhumby, numele companiei de servicii de gestionare a datelor clienţilor (*customer data servicest* engl.) înfiinţate de soţii Edwina Dunn şi Clive Humby în 1989, în propria bucătărie, care a ajuns în prezent să aibă sucursale în peste 30 de ţări şi să fie estimată a avea o valoare cuprinsă între unu şi trei miliarde de dolari americani (n. red.).

** „Preferinţele mele” (engl.) (n.tr.).

Nu toţi, însă. La puţin timp după lansarea acestui element, o femeie a contactat compania Tesco şi s-a plâns că informaţiile care i se ofereau ei erau greşite. Făcuse cumpărături on-line şi, printre produsele ei din lista *My Favourites* observase şi prezervative! Nu puteau fi ale soţului ei, explicase clienta, pentru că el nu folosea aşa ceva. La această solicitare, analiştii de la Tesco au studiat informaţiile şi au descoperit că lista era corectă. Cu toate acestea, în loc să fie cauza unei rupturi conjugale, au luat decizia diplomatică de a-şi cere scuze pentru „informaţiile incorecte” şi de a şterge obiectele ofensatoare din lista de produse favorite a clientei.

După cum a scris Clive Humby în cartea lui despre Tesco, procedura a devenit acum politica neoficială a companiei. Astfel, ori de câte ori în datele informatizate ale unui client apare ceva care dezvăluie cam prea multe, cei de la Tesco îşi cer scuze şi şterg informaţiile. Este o atitudine pe care şi-a însuşit-o şi Eric Schmidt, care, pe când era preşedintele executiv al Google, spunea că încerca să gândească în termenii unei limite imaginare care ar despărţi lucrurile acceptabile de acelea sinistre (*creepy line*):

Politica Google este de a se apropia cât mai mult de graniţa care ne desparte de chestiile de-a dreptul sinistre, fără să o depăşească, însă.4

Totuşi, dacă vei culege suficiente informaţii, e greu de spus ce vei descoperi. Deoarece cumpărăturile nu reprezintă numai ceea ce consumi. Ele sunt ceva profund

personal. Dacă vei studia cu atenţie comportamentul la shopping al cuiva, el îţi va dezvălui, de cele mai multe ori, tot felul de amănunte despre cine este acesta ca persoană. Uneori – cum a fost în cazul prezervativelor –, va divulga lucruri pe care, poate, ar fi mai bine să nu le ştii. Însă, cel mai adesea, în tot acest noian de informaţii, se ascund frânturile unor descoperiri uimitoare care pot fi folosite în beneficiul unei companii.

MAGAZINUL TARGET

În anul 2002, Target, compania de supermarketuri americane specializate în comercializarea produselor cu reduceri, a început să caute patternuri mai neobişnuite în informaţiile colectate.5 Target vinde orice, de la lapte şi banane, la jucării de pluş sau mobilă de grădină, şi – aproape cam ca oricare alt comerciant cu amănuntul de la începutul secolului al XXI-lea – are modalităţile lui de a utiliza cifrele de pe cardurile de credit şi răspunsurile la sondaje pentru a face legătura între clienţi şi tot ceea ce au achiziţionat aceştia vreodată din magazin, ceea ce le permite să analizeze ce cumpără oamenii.

O ştire (care nu mai trebuie să le fie repetată cititorilor americani) a făcut de ruşine compania în Statele Unite: se pare că, la un moment dat, analiştii de la Target şi-au dat seama că o creştere bruscă a achiziţionării de către femei a loţiunilor de corp neparfumate indica, adesea, că respectivele cliente urmau să se înscrie în registrul magazinului[16] pentru ca rudele şi prietenii lor să poată cumpăra articolele pe care şi-ar fi dorit să le primească în dar la *baby-shower*[17] în datele informatizate fusese descoperit un indiciu plin de semnificaţii: când

16 în SUA chiar şi supermarketurile le permit clienţilor să îşi creeze o listă cu lucrurile aflate la vânzare în respectiva unitate comercială pe care ar dori să le primească drept cadou cu diverse ocazii (nuntă, botezul copiilor etc.). O astfel de listă poartă numele de gift registry (engl., „registru al cadourilor”) (n. red.).

17 Petrecere organizată în timpul sarcinii de către viitoarea mamă sau de către prietenii ori rudele acesteia, pentru a celebra apropiata sosire pe lume a copilului (n. tr.).

femeile intrau în al doilea trimestru de sarcină şi începeau să îşi facă griji din cauza vergeturilor, cumpărarea de creme hidratante pentru menţinerea supleţei pielii era un indiciu care prevestea ceea ce avea să urmeze. Dacă ai fi derulat înapoi istoricul cumpărăturilor, ai fi observat că aceleaşi femei trecuseră pe la Target şi se aprovizionaseră cu diverse vitamine şi suplimente alimentare, cum ar fi calciu şi zinc. Dac-ai fi derulat apoi înainte aceeaşi listă de cumpărături, informaţiile ţi-ar fi sugerat chiar şi data la care se născuse bebeluşul – indicată de faptul că după aceea, femeia cumpăra din magazin pungi foarte mari de vată.6

Mamele însărcinate sunt visul oricărui comerciant cu amănuntul. Dacă ţi-ai asigurat fidelitatea lor cât timp erau însărcinate, sunt şanse mari ca ele să continue să-ţi folosească produsele mult timp după naşterea copilului. În definitiv, obiceiurile la shopping se formează iute atunci când un bebeluş flămând care urlă din toţi rărunchii îţi solicită atenţia în timpul cumpărăturilor săptămânale. Informaţiile de acest gen erau de o inestimabilă valoare, căci îi puteau oferi companiei Target un avantaj faţă de rivalele sale în atragerea cumpărăturilor pe care urma să le facă mama respectivă.

De aici încolo, lucrurile au fost simple. Target a rulat un algoritm care avea să ofere un punctaj clientelor, în funcţie de probabilitatea ca acestea să fie însărcinate. Dacă probabilitatea depăşea un anumit prag, magazinul avea că îi trimită în mod automat femeii în cauză o serie de cupoane cu reduceri pentru tot soiul de lucruri pe care aceasta le-ar fi putut considera utile: scutece, loţiuni, şerveţele umede pentru copii şi aşa mai departe.

Amintita practică era, pe atunci, necontroversată. Apoi, cam la un an după introducerea acestui instrument, tatăl unei adolescente a intrat ca o furtună într-un magazin Target din Minneapolis, cerând să discute cu responsabilul unităţii. Fiicei sale îi fuseseră trimise cupoane cu produse pentru femei gravide, iar el era revoltat că magazinul

părea să considere fireşti sarcinile în rândul adolescentelor. Responsabilul şi-a cerut mii de scuze, iar câteva zile mai târziu a sunat la domiciliul bărbatului pentru a-şi exprima încă o dată părerea de rău pentru toată încurcătura. Numai că, între timp, conform relatării din *New York Times*, tatăl s-a trezit, la rândul său, că avea de prezentat nişte scuze.

„Am avut o discuţie cu fiica mea”, i-a spus el responsabilului. „Se pare că în casa mea s-au petrecut unele lucruri de care nu prea am avut habar. Fiica mea urmează să nască în august”. Nu ştiu ce părere aveţi, dar un algoritm care va înştiinţa un părinte că fiica sa este însărcinată înainte ca el să aibă prilejul să afle personal acest lucru, mie mi se pare că trece mult dincolo de graniţa care ne desparte de chestiile de-a dreptul sinistre. Această situaţie stânjenitoare nu a fost, însă, suficientă pentru a convinge firma să renunţe definitiv la respectiva strategie de marketing.

Un director al companiei Target a explicat următoarele:

Am aflat că, atât timp cât o femeie însărcinată nu crede că a fost urmărită, ea va folosi cupoanele noastre de reduceri.

Va presupune numai că toţi cei din blocul în care locuieşte au primit acelaşi pliant pentru scutece şi pătuţuri. Câtă vreme nu o speriem, metoda funcţionează.

Aşadar, Target încă are un algoritm capabil să anticipeze sarcinile care rulează în fundal – aşa cum procedează, în prezent, majoritatea comercianţilor. Singura diferenţă este că, în prezent, când va trimite cupoane cu reduceri pentru diversele produse dedicate femeilor însărcinate, le va amesteca printre altele pentru articole de uz general, astfel încât clientele să nu observe că au fost vizate în mod direct. O reclamă la pătuţuri pentru nou-născuţi ar putea să apară lângă una pentru pahare de vin. Sau un cupon cu reduceri pentru hăinuţe destinate bebeluşilor va fi postată pe site-ul companiei

lângă o reclamă la o marcă de apă de colonie.

Target nu este singura firmă care foloseşte aceste metode. Relatările despre ceea ce se poate deduce din datele personale ale cuiva ajung rareori în presă, însă algoritmii sunt prezenţi, ascunzându-se, tăcuţi, în spatele faţadelor corporatiste. Cu aproximativ un an în urmă, am stat de vorbă cu analistul-şef[18] al unei companii care vindea asigurări. Prin intermediul unei strategii de marketing de loializare[19] a clienţilor unui supermarket, el şi colegii lui aveau acces la toate detaliile privind obiceiurile de shopping ale acestora. Analizând informaţiile, analiştii au descoperit că toţi cei care găteau acasă erau mai puţin înclinaţi să emită cereri de despăgubire, deşi locuinţele le erau asigurate, fiind, prin urmare, clienţi mai profitabili pentru compania sa. Acesta este un exemplu de descoperire care pare de bun-simţ, în mod intuitiv. Probabil că nu există prea multe puncte comune între cei care sunt dispuşi să investească timp, efort şi bani pentru a pregăti de la zero diverse feluri de mâncare complicate şi cei care îşi lasă copiii să joace fotbal în casă. Dar de unde au ştiut analiştii companiei de asigurări care cumpărători găteau acasă şi care nu? Ei bine, pur şi simplu, prezenţa unor anumite articole în coşul de cumpărături al unor persoane corespundea cu o rată redusă de cereri de despăgubire. Cel mai important articol, mi-a declarat analistul-şef, acel produs care dezvăluie mai mult decât orice altceva că respectivul cumpărător este o persoană responsabilă şi mândră de locuinţa sa, e… feniculul proaspăt.

---

18 Chief data officer (engl., CDO), director al unei corporaţii a cărui responsabilitate o constituie gestionarea şi utilizarea datelor informatizate prin diverse mijloace: colectare, prelucrare, analizare, data mining, schimb de informaţii ş.a. (n. red.).

19 Loyalty scheme (engl.), strategie de marketing care îşi propune să îi transforme pe clienţii deja existenţi în cumpărători dedicaţi, care nu mai folosesc alte magazine. Deşi termenul nu este atestat de dicţionarele româneşti, am optat să îl păstrăm deoarece este larg folosit în lucrările de marketing, alături de „fidelizare” (n. red.).

Dacă poţi să deduci una ca asta din comportamentul la shopping al persoanelor în lumea reală, imaginaţi-vă ce se poate deduce dacă ai acces la mai multe informaţii. Imaginaţi-vă numai cât de multe lucruri se pot afla despre cineva dacă dispui de evidenţa întregii sale activităţi on-line!

VESTUL SĂLBATIC

Palantir Technologies este una dintre cele mai de succes companii abia înfiinţate din Sillicon Valley. A fost întemeiată în 2003 de Peter Thiel (ultracelebrul cofondator al PayPal), iar, la cea mai recentă evaluare disponibilă înainte de publicarea acestei cărţi, s-a estimat că ar valora uluitoarea sumă de 20 de miliarde de dolari.7 Are cam aceeaşi valoare de piaţă ca şi Twitter, deşi sunt şanse să nu fi auzit niciodată de acei de la Palantir. Totuşi – şi vă rog să mă credeţi când vă spun asta!

— Palantir a auzit, mai mult ca sigur, de fiecare dintre noi.

Palantir este face parte dintr-o nouă specie de companii, cunoscute drept „brokeri de date”[20], care cumpără şi colectează informaţiile personale ale oamenilor, pentru ca apoi să le revândă sau să le împartă cu alte firme pentru a obţine profit. Şi există multe alte companii asemănătoare: Acxiom, Corelogic, Datalogix, ebureau – o grămadă de firme uriaşe cu care probabil că nu aţi interacţionat niciodată în mod direct, dar care vă monitorizează şi vă analizează neîncetat comportamentul.8

De fiecare dată când faceţi cumpărături on-line, de fiecare dată când vă înscrieţi la un newsletter, vă înregistraţi pe un site, vă interesaţi de o maşină nouă, completaţi un certificat de garanţie, cumpăraţiolocuinţănouăsauvăânregistraţi să votaţiadică de fiecare dată când oferiţi cuiva orice fel de date personale –, toate aceste informaţii sunt adunate şi vândute apoi unui broker de date. Vă aduceţi aminte când i-aţi descris unui agent imobiliar ce fel de locuinţă căutaţi? Ei bine,

20 Data broker (engl.) (pentru definiţie, vezi glosar, p. 316) (n. red.).

informaţiile acelea i-au fost deja vândute unui broker de date. Mai ştiţi toate amănuntele acelea pe care le-aţi tastat la un moment dat pe un website care compara preţurile diferitelor poliţe de asigurări? Şi ele i-au fost deja vândute unui broker de date. În anumite situaţii, chiar şi întregul istoric al browserului vostru poate fi compactat şi vândut.9

Cade în sarcina brokerului să combine toate aceste informaţii, să coreleze diferitele fragmente din toate frânturile digitale pe care le-a cumpărat şi le-a strâns, iar apoi să creeze un singur fişier detaliat despre fiecare dintre noi: un profil psihologic alcătuit din date, urma ta digitală, „umbra" ta informatică. În sensul cel mai literal, în unele dintre aceste baze de date ale brokerilor de informaţii ai putea să deschizi un fişier digital purtând numărul tău de identificare (un număr care ţie nu-ţi va fi comunicat niciodată) care conţine urmele a tot ceea ce ai făcut vreodată: nume, data naşterii, credinţă religioasă, obiceiuri de vacanţă, mod de utilizare a cardului de credit, avere, greutate, înălţime, afinităţi politice, ce jocuri obişnuieşti să joci, ce dizabilităţi ai, ce medicamente foloseşti, dacă ai făcut vreodată vreun avort, dacă părinţii tăi sunt divorţaţi sau nu, dacă ai o predispoziţie spre a deveni dependent cu uşurinţă, dacă ai fost victima unui viol, ce părere ai despre interzicerea armelor de foc, care este orientarea sexuală pe care o afişezi şi care e adevărata ta orientare sexuală sau cât de credul eşti… Practic, pentru fiecare dintre noi există mii şi mii de detalii, incluse în mii şi mii de categorii, în fişiere stocate pe nişte servere aflate cine ştie pe unde.10

Asemenea predicţiilor făcute de analiştii companiei Target despre sarcinile clientelor, multe dintre aceste informaţii pot fi deduse cu uşurinţă. Un abonament la revista *Wired* ar putea însemna că eşti interesat de tehnologie; un permis de portarmă – că te pasionează vânătoarea. Pe tot acest parcurs, brokerii folosesc algoritmi inteligenţi, dar simpli, pentru a spori mereu cantitatea de informaţii pe care o deţin. Este exact ceea ce

făceau şi supermarketurile, dar pe o scară cu mult mai mare.

Şi, datorită acestor profiluri, avem parte şi de o mulţime de beneficii. Astfel, brokerii de date îşi folosesc toate cunoştinţele pe care le deţin despre cine suntem şi cum ne comportăm on-line pentru a-i împiedica pe escroci să se dea drept consumatori nevinovaţi. De asemenea, tot pentru că ni se cunosc preferinţele, dar şi lucrurile care ne displac, reclamele care ne sunt prezentate în timp ce navigăm pe internet sunt pe cât de relevante e cu putinţă pentru interesele şi nevoile noastre. Aceasta aproape că ne asigură o experienţă mai plăcută decât aceea în care am fi agresaţi zi de zi de reclame transmise la grămadă, tuturor,[21] la vreo asigurare de credit oarecare ori la serviciile oferite de cine ştie ce avocat specializat în procesele deschise de aceia care pretind c-ar fi fost răniţi de diverse persoane. Mai mult decât atât, pentru că mesajele pot fi ţintite direct spre consumatorii potriviţi, care sunt interesaţi de ele, publicitatea devine mai ieftină, în general, astfel încât firmele mici care oferă produse de excepţie pot ajunge la noi nişe de piaţă, la noi categorii de consumatori, ceea ce este bine pentru toată lumea.

Însă, aşa cum sunt convinsă că v-aţi dat seama deja, pot apărea şi o mulţime de probleme în momentul în care începi să ne separi pe noi, ca indivizi, într-o serie de categorii. Voi trece la acest subiect în curând, dar, mai întâi, cred că merită să explic, pe scurt, procesul nevăzut care are loc pentru ca o reclamă on-line să ajungă la noi atunci când dăm clicuri pe diverse site-uri de pe net, precum şi rolul pe care-l joacă brokerii de date în acest proces.

Să ne imaginăm, aşadar, că aş deţine o agenţie de voiaje de lux cu numele fictiv de Fry’s. În decursul anilor, am reuşit să fac oamenii să se înregistreze pe site-ul

21 Mass market adverts (engl.), „[campanie de] reclame care face parte dintr-o strategie de tip mass market” (vezi glosar, p. 318) (n. red.).

companiei mele dacă au manifestat interes faţă de el, iar de acum am o listă cu adresele lor de e-mail. Dacă aş vrea să aflu mai multe despre utilizatorii mei – de exemplu, ce fel de vacanţe preferă –, aş putea să trimit lista cu adresele lor de e-mail unui broker de date, care le-ar căuta numele în sistem şi mi-ar înapoia lista, completată însă cu diverse informaţii relevante. Este ca şi când s-ar adăuga o coloană nouă într-un tabel. De acum, ori de câte ori un utilizator va vizita site-ul Fry's, eu voi vedea că respectivul are o anumită preferinţă pentru insulele tropicale şi, astfel, o să-i ofer o reclamă la o excursie în Hawaii.

Aceasta ar fi prima mea opţiune. Pentru cea de-a doua, haideţi să ne imaginăm că Fry's are ceva spaţiu suplimentar pe site, pe care sunt dispusă să-l vând altor agenţii de publicitate. Voi contacta din nou un broker de date şi-i voi da informaţiile pe care le deţin despre utilizatorii mei. Brokerul va căuta companii care să-şi dorească să-şi prezinte reclamele pe site-ul agenţiei mele. Şi, doar de dragul de a continua această povestioară, haideţi să ne imaginăm că o companie care vinde o loţiune de protecţie solară şi bronzare este deosebit de interesată să-i facă reclamă. Ca să o conving că Fry's deţine audienţa căreia vânzătorul de cosmetice respectiv ar dori să i se adreseze, brokerul ar putea să-i dezvăluie unele trăsături ale utilizatorilor site-ului meu pe care le-a dedus analizându-le datele informatizate: spre exemplu, care este procentul utilizatorilor mei care au păr roşcat ori ceva de genul acesta. Sau vânzătorul loţiunii ar putea să-i ofere propria listă cu adresele de e-mail ale utilizatorilor săi, iar brokerul şi-ar putea da seama exact câte puncte comune există între cele două segmente de public. Dacă vânzătorul loţiunii de protecţie solară este de acord, reclamele lui vor apărea pe site-ul Fry's – şi atât brokerul, cât şi eu, vom fi plătiţi.

Până acum, metodele acestea nu depăşesc cu mult strategiile pe care specialiştii în marketing le-au utilizat dintotdeauna pentru a-şi viza clienţii. Dar deja, când

trecem la cea de-a treia opţiune, lucrurile încep să devină cam prea sinistre. De data aceasta, Fry’s caută nişte clienţi noi şi speră să atragă bărbaţi şi femei care au peste 65 de ani, cărora le plac insulele tropicale şi dispun de venituri mari, nădăjduind că aceştia vor dori să meargă în vreuna dintre noile croaziere de lux în Caraibe pe care le oferă. Aşa că, în calitate de patron, voi contacta un broker de informaţii, care va căuta în baza lui de date şi-mi va găsi o listă cu diverse persoane care se potrivesc acestei descrieri.

Să ne imaginăm, aşadar, că numele dumneavoastră se află pe lista aceea. Brokerul nu-i va dezvălui, în niciun caz, vreo identitate companiei Fry’s. Însă va şti ce alte site-uri vizitaţi cu regularitate. Sunt şanse ca şi brokerul să aibă o relaţie de afaceri cu unul dintre deţinătorii site-urile voastre preferate. Poate că e vorba de un site de socializare, poate că e vorba de unul de ştiri – ceva de genul acesta. Imediat ce v-aţi autentificat, fără să bănuiţi nimic, pe unul dintre site-urile preferate, brokerul va primi o înştiinţare (*ping*) care îl anunţă că navigaţi pe una dintre paginile acestuia. Practic instantaneu, brokerul va răspunde prin plasarea unui „steguleţ” – cunoscut drept cookie în calculatorul tău. Această *cookie** acţionează ca un semnal pentru toate tipurile de site-uri de pe internet, transmiţându-le că sunteţi o persoană căreia ar trebui să i se prezinte o reclamă la croazierele în Caraibe ale companiei Fry’s. Aşa că, fie că vă doriţi acest lucru, fie că nu, reclamele vă vor urmări pretutindeni, oriunde veţi naviga pe internet.

— Reclamele nu sunt singurul motiv pentru care există cookieuri, căci acestea mai sunt folosite de site-uri ca să vadă dacă sunteţi autentificat sau nu (şi să ştie dacă este sigur să trimită pe această cale informaţii cu caracter confidenţial) şi ca să vadă dacă sunteţi un vizitator care revine pe o anumită pagină (pentru a putea obţine o creştere a preţurilor pe site-ul unei companii aeriene, de exemplu, sau pentru a vă trimite prin e-mail un cod pentru

obţinerea unei reduceri la un magazin virtual de îmbrăcăminte) (n.a.).

Şi aici apare prima problemă. Ce se întâmplă dacă nu-ţi doreşti să vezi acea reclamă? Este adevărat, să fii bombardat cu imagini ale unor croaziere în Caraibe nu poate fi decât un inconvenient minor, dar există şi reclame care chiar pot afecta profund o persoană.

Când Heidi Waterhouse a pierdut o sarcină pe care şi-o dorise nespus de mult11, ea s-a dezabonat de la toate e-mailurile săptămânale care o ţineau la curent cu dezvoltarea bebeluşului, spunându-i de mărimea cărui fruct ajunsese fătul într-o anumită etapă. Şi-a anulat, totodată, şi abonamentele la toate listele de corespondenţă, precum şi toate listele de lucruri dorite[22] la care se înscrisese în aşteptarea plină de înfrigurare a momentului naşterii. Numai că, aşa cum avea să le relateze ea unui grup de dezvoltatori la o conferinţă din 2018, nicio putere din lume nu o putea ajuta să îşi anuleze „abonarea” la puzderia de reclame pentru femeile însărcinate care o urmăreau pretutindeni pe internet. Urmele digitale ale sarcinii au continuat să circule singure, ca nişte umbre informatice, deşi, de acum, nu mai existau nicio mamă şi niciun bebeluş. „Niciunul dintre cei care au creat sistemul nu s-a gândit vreodată la astfel de consecinţe”, a explicat ea.

Fiindcă este un sistem care, fie din cauza unei omisiuni necugetate, fie din cauza proiectării intenţionate, are potenţialul de a deveni abuziv. Cămătarii care dau credite „doar cu buletinul” îl pot utiliza ca să vizeze în mod direct persoanele cărora nu le-ar fi acordate niciodată împrumuturi avantajoase fiindcă nu întrunesc condiţiile; reclamele la pariuri ar putea fi direcţionate către

22 Wish list (engl.), listă pe care o persoană o alcătuieşte pentru a-i ajuta pe cei care ar dori să îi facă daruri să aleagă obiecte pe care chiar şi le doreşte. în anumite ţări există obiceiul ca oamenii să facă schimb de astfel de liste cu diverse ocazii: Crăciun, nuntă, botez etc. Un gift registry (vezi prima notă de la p. 48) este un caz special de listă cu lucruri dorite (n. red.).

persoanele care frecventează site-urile cu jocuri de noroc… Şi există chiar temeri că acest gen de elaborare a profilului cu ajutorul informaţiilor poate fi folosit chiar şi împotriva populaţiei: astfel, cei pasionaţi de motociclete pot fi consideraţi a avea un hobby periculos, iar persoanele care consumă dulciuri fără zahăr pot fi semnalate ca fiind diabetice şi tuturor li se poate refuza încheierea unei asigurări din aceste cauze.

Un studiu din 2015 a arătat că Google afişa mult mai puţine reclame la posturi de conducere bine plătite femeilor care navigau pe internet, decât bărbaţilor.12

Iar după ce o profesoară afro-americană de la Harvard a descoperit că, atunci când îşi căuta numele pe Google, primea după aceea reclame care vizau persoane cu antecedente penale (şi, drept urmare, a fost nevoită să-i demonstreze unui posibil angajator că nu avusese niciodată probleme cu poliţia), aceasta a început să cerceteze reclamele primite de diferite grupuri etnice. Din câte a descoperit, căutările pentru „nume cu rezonanţă africană" aveau şanse disproporţionat de mari să fie însoţite de reclame care conţineau cuvântul „arest" (de exemplu: „Aţi fost arestat?"), decât acelea pentru „numele care păreau să le aparţină unor albi".13

Aceste metode nu sunt folosite numai de către brokerii de date, căci există diferenţe foarte mici între modul în care lucrează cei din urmă şi acela în care operează Google, Facebook, Instagram şi Twitter. Aceşti giganţi ai internetului nu fac bani pentru că atrag utilizatori, aşa că modelele lor de afaceri se bazează pe ideea de marketing personalizat *Cmicrotargeting*). Toate aceste platforme sunt, de fapt, nişte motoare uriaşe de furnizare de reclame, câştigând bani din faptul că milioane de utilizatori, angajaţi activ pe site-urile lor, dau clicuri în stânga şi în dreapta, citesc postări sponsorizate, urmăresc videoclipuri sponsorizate, privesc fotografii sponsorizate.

În orice colţişor de internet pe care-l veţi accesa vor sta ascunşi în fundal nişte algoritmi care comercializează

informaţii despre dumneavoastră pe care nici nu ştiaţi că le deţine cineva şi pe care nu le-aţi oferit nicicând nimănui de bunăvoie. Aceşti algoritmi ne-au transformat secretele cele mai personale şi mai intime într-o marfa.

Din nefericire, în numeroase ţări legea nu face prea mult ca să te protejeze. Activitatea brokerilor de date este, în marea sa majoritate, nereglementată şi – mai ales în America – ocaziile de a le mai reduce puterea pe care o deţin au fost ignorate în mod repetat de către guvern. În martie 2017, de exemplu, Senatul Statelor Unite a votat pentru eliminarea reglementărilor care i-ar fi împiedicat pe brokerii de date să vă vândă istoricul navigării pe internet în lipsa consimţământului dumneavoastră explicit. Aceste reglementări fuseseră, în prealabil, aprobate în octombrie 2016 de către Comisia Federală de Comunicaţii, însă, după schimbarea care a avut loc în guvern la sfârşitul acelui an[23], au fost respinse de noua majoritate republicană a FCC (Federal Communications Commission) [24] şi de republicani, în Congres.14

Însă ce înseamnă, mai exact, toate acestea pentru confidenţialitatea dumneavoastră? Ei bine, daţi-mi voie să vă relatez o investigaţie condusă de jurnalista germană Svea Eckert şi analistul de date Andreas Dewes, care ar trebui să vă ofere o imagine mai clară.15

Eckert şi echipa ei au inventat un fals broker de date şi l-au folosit pentru a cumpăra datele referitoare la navigarea anonimă pe net a trei milioane de cetăţeni germani. (A fost uşor să intre în posesia istoricului navigării pe internet a acestora. O mulţime de companii aveau de vânzare o groază de informaţii de acest gen despre clienţi britanici sau americani – singura provocare a fost găsirea datelor care aveau legătură cu Germania.) Informaţiile în sine fuseseră culese de un *plug-in* al Google

23 La 8 noiembrie 2016 au avut loc în SUA alegerile prezidenţiale al căror câştigător a fost Donald Trump. (n. tr.).

24 Comisia Federală de Comunicaţii, agenţie independentă a guvernului american însărcinată cu reglementarea telecomunicaţiilor din statele americane (n. red.).

Chrome pe care utilizatorii îl descărcaseră de bunăvoie, în completă necunoştinţă de cauză că acesta îi mai şi spiona pe parcurs.*

În total, s-a obţinut o listă imensă de URL-uri. O înregistrare a tot ceea ce respectivele persoane văzuseră on-line în decursul unei luni. Fiecare căutare, fiecare pagină, fiecare clic. Totul, pus în mod legal la vânzare.

Pentru Svea Eckert şi colegii ei, singura problemă a fost că informaţiile de navigare erau anonime. Ceea ce ar fi trebuit să fie nişte veşti bune pentru toţi oamenii al căror istoric de browsing fusese vândut, nu-i aşa? Căci, în cazul acesta, niciunul dintre ei n-ar mai fi trebuit să se înroşească de ruşine, corect? Nu, nu! Greşit! Aşa cum a explicat echipa de investigaţii într-o prezentare la DEFCON în anul 2017, dezanonimizarea imenselor baze de date cu istoricele de navigare pe internet a fost fenomenal de simplă.

Iată cum au reuşit s-o facă. Uneori, au existat indicii directe despre identitatea persoanei chiar în URL-uri. De exemplu, toţi cei care au vizitat Xing.com, echivalentul german al Linkedin. Dacă dai clic pe fotografia de profil a website-ului Xing, eşti trimis la o pagină cu o adresă de genul:

www.xing.com/profile/Hannah Fry7 scomxbp

Numele te trădează imediat, iar textul care urmează după numele de utilizator înseamnă că utilizatorul este logat şi îşi vizionează propriul profil, astfel că echipa putea fi convinsă că persoana respectivă îşi privea pagina personală. Acelaşi lucru este valabil şi în cazul lui Twitter. Oricine îşi verifica propria pagină de Twitter Analytics

— Acel *plug-in*, ironic denumit *The Web of Trust* („Reţeaua încrederii”), stipula asta în mod clar, negru pe alb, ca parte a termenilor şi condiţiilor (n.a.).

se dădea de gol prin această acţiune. Pentru cei în cazul cărora în setul de date deţinut nu exista şi un identificator care să poată fi detectat instantaneu, echipa avea un alt as în mânecă… Toţi cei care postaseră vreodată

un link Online (poate printr-un tweet despre un site sau prin partajarea playlist-ului de pe YouTube), practic, oricine lăsase în public o urmă digitală, o» umbră" informatică legată de numele său adevărat, a fost demascat în mod accidental prin această acţiune. Echipa de cercetători a folosit un algoritm simplu pentru a pune în corespondenţă persoanele reale cu cele anonime16, filtrându-le lista de URL-uri pentru a găsi pe cineva din setul de date care vizitase aceleaşi site-uri la aceleaşi date calendaristice şi la aceleaşi repere orare la care fuseseră postate linkurile on-line. În cele din urmă, au obţinut numele şi prenumele tuturor persoanelor din setul de date, precum şi accesul complet la istoricul de navigare dintr-o lună de zile a milioane de germani.

Printre aceste trei milioane de oameni se numărau şi câteva persoane importante: un politician care căutase nişte medicamente on-line. Un ofiţer de poliţie care copiase un document dintr-un dosar delicat şi-l lipise într-o fereastră din Google Translate, iar detaliile acestuia apăruseră apoi în URL, putând fi văzute de cercetători. Şi un judecător al cărui istoric de căutare dezvăluia vizitarea zilnică a unui domeniu destul de „specializat" de pe internet. Iată, în cele ce urmează, o mică selecţie a site-urilor pe care le vizitase acesta într-un interval de opt minute, în august 2016:

18.22: http://www.tubegalore.com/video/amature-povex-wife-in-leather-pants-gets-creampie42945.html

18.23: http://www.xxkingtube.com/video/povwifeyon sex stoolwithbeadedthonggetscreampie4814.html 18.24: http://dejdiamster.com/movies/924590/officelady impanstrubbingridinghestofanlife.html

18.27: http://www.tubegalore.com/youngtube/5762-l / pageO

18.30: http://www.keezmovies.com/video/sexydominatrix-milks-him-dry-1007114? utmsources

Pe lângă aceste sesiuni zilnice de navigare pe

internet, judecătorul mai căuta, cu regularitate, on-line nume de nou-născuţi, cărucioare şi spitale cu maternitate. Echipa a tras concluzia că în perioada aceea partenera sa aştepta un copil.

Să lămurim, însă, un lucru: judecătorul nu făcea nimic ilegal. Mulţi oameni – printre care mă număr şi eu – vor susţine că nu făcea nimic rău. Însă nu este mai puţin adevărat că aceste informaţii ar fi fost utile pentru cineva care ar fi vrut să-l şantajeze sau să-i pună familia într-o situaţie stânjenitoare.

Iar acesta este locul de unde începem să ne îndepărtăm şi să depăşim graniţa care ne desparte de lucrurile sinistre – atunci când informaţii private şi delicate despre propria persoană, culese fără ştirea ta, sunt folosite ulterior pentru a te manipula. Ceea ce, evident, este exact ce a făcut firma britanică de consultanţă politică Cambridge Analytica.

CAMBRIDGE ANALYTICA

Probabil că deja cunoaşteţi cea mai mare parte a poveştii, începând din anii 1980, psihologii folosesc un sistem de cinci caracteristici pentru cuantificarea personalităţii unui individ. Primeşti un punctaj pentru fiecare dintre următoarele trăsături: deschiderea (*openness*) faţă de noile experienţe, conştiinciozitate (*conscientiousness*), caracter extravertit (iextraversion), caracter agreabil (*agreeableness*) şi caracter nevrotic (*neuroticism*).[25] Luate împreună, acestea oferă o metodă standard deosebit de utilă pentru a descrie ce tip de persoană eşti.

25 Cunoscut în psihologie drept modelul celor cinci trăsături sau modelul celor cinci factori (five-factor model FFM, engl.) ori modelul OCEAN - acronim al denumirilor în engleză ale celor cinci trăsături ale personalităţii: openness (engl., „deschidere”), conscientiousness (engl., „conştiinciozitate”), extraversión (engl., „caracter extravertit”), agreeableness (engl., „caracter agreabil) şi neuroticism (engl., „caracter nevrotic”) -, reprezintă un model de descriere a spectrului personalităţilor propus în 1980 de psihologii Lewis Goldberg, Naomi Takemoto-Chock, Andrew Comrey şi John M. Digman şi rafinat de atunci de mai mulţi cercetători (n. red.).

În 2012, cu un an înainte de intrarea în scenă a Cambridge Analytica, mai mulţi oameni de ştiinţă de la Universitatea Cambridge şi de la Universitatea Stanford au început să caute o legătură între cele cinci trăsături de personalitate şi paginile cărora oamenii le dădeau like-uri pe Facebook.17 Au construit un chestionar pe Facebook în acest scop, lăsând utilizatorii să efectueze teste reale de psihometrie în speranţa că vor descoperi o legătură între caracterul autentic al unei persoane şi identitatea sa on-line. Cei care au descărcat chestionarul le-au oferit, cu bună ştiinţă, informaţii şi despre istoricul like-urilor personale de pe Facebook, iar, prin intermediul răspunsurilor la o serie de întrebări, valorile personalităţii lor reale, procentele în care deţineau cele cinci caracteristici.

Este uşor să ne imaginăm cum ar putea fi corelate like-urile date cu personalitatea. Aşa cum a subliniat echipa de cercetători în lucrarea pe care a publicat-o în anul următor18, aceia cărora le plac Salvador Dalí, meditaţia sau conferinţele de pe platforma TED vor obţine aproape sigur punctaje mari la deschiderea către noi experienţe. În schimb, persoanele cărora le plac petrecerile, dansul şi personajul Snooki din serialul de televiziune *Jersey Shore* tind să fie ceva mai extravertite. Cercetarea a fost încununată de succes. După ce a stabilit această legătură, echipa a construit un algoritm care putea deduce personalitatea cuiva numai din like-urile pe care le dădea pe Facebook.

Până să le apară cel de-al doilea studiu, în 201419, cercetătorii susţineau deja că, dacă reuşeai să strângi 300 de like-uri din profilul de pe Facebook al unei persoane, algoritmul putea să-i analizeze acesteia caracterul mai precis decât soţul sau soţia sa.

Revenind în prezent, trebuie spus că acest grup dedicat cercetării ştiinţifice, mai precis, Centrul de Psihometrie al Universităţii Cambridge, a extins algoritmul pentru a face predicţii despre personalitatea cuiva şi cu

ajutorul mesajelor postate de acesta pe Twitter. Există un site, accesibil oricui, unde poţi să încerci chiar tu algoritmul. Din moment ce profilul meu pe Twitter este public, m-am gândit să încerc predicţiile cercetătorilor asupra propriei persoane, aşa că le-am trimis istoricul mesajelor de pe Twitter şi, pentru a avea un termen de comparaţie, am completat un studiu de personalitate bazat pe un chestionar tradiţional. Algoritmul a reuşit să-mi evalueze corect trei însuşiri din cinci, cu toate că, din câte s-a dovedit din studiul tradiţional de personalitate, sunt mult mai extravertită şi nevrotică decât lasă profilul meu de pe Twitter să se înţeleagă.*

Toate aceste activităţi au fost motivate de faptul că ele puteau fi de folos în publicitate. Astfel, în 201720, aceeaşi echipă de cadre universitare a continuat cu un alt experiment, trimiţând reclame adaptate la trăsăturile de personalitate ale unei persoane. Folosind platforma Facebook, echipa a furnizat reclame la produse de înfrumuseţare persoanelor extravertite, folosind sloganul: „Dansează ca şi când nu te-ar vedea nimeni (deşi te vede toată lumea)”, pe când persoanele introvertite au văzut imaginea unei fete care îşi zâmbea, privindu-se în

— Această combinaţie pare să sugereze că aş posta mai multe lucruri dacă nu mi-aş face atâtea griji în legătură cu felul în care vor fi primite de ceilalţi (n.a.).

oglindă, însoţită de textul: „Frumuseţea nu trebuie neapărat să fie stridentă”.

Într-un experiment desfăşurat în paralel, persoanelor din grupul-ţintă care erau într-o mare măsură deschise spre acceptarea experienţelor noi li s-au arătat reclame la cuvinte încrucişate folosind o imagine cu textul: „Aristotel? Insulele Seychelles? Daţi-vă frâu liber creativităţii şi stimulaţi-vă imaginaţia cu o cantitate nelimitată de cuvinte încrucişate!” Aceleaşi cuvinte încrucişate erau apoi prezentate pe ecran persoanelor care erau doar într-o măsură redusă deschise spre experienţe noi, dar folosindu-se îndemnul: „Instalaţi-vă comod alături de jocul preferat

dintotdeauna! Cuvintele încrucişate care au constituit o provocare pentru generaţii întregi de pasionaţi”. Per ansamblu, echipa a susţinut că stabilirea unei corelaţii între reclame şi caracterul unei persoane a generat cu până la 40% mai multe clicuri şi cu până la 50% mai multe achiziţii decât folosirea reclamelor generice, nepersonalizate. Iar pentru orice agenţie de publicitate, acestea sunt rezultate de-a dreptul impresionante!

În tot acest timp, pe când cadrele universitare îşi publicau lucrarea, alţii le implementau metodele. Printre ei s-a numărat şi compania Cambridge Analytica în perioada în care a lucrat pentru campania electorală a lui Donald Trump.

Haideţi, acum, să ne întoarcem puţin în timp! Există puţine dubii că, de fapt, Cambridge Analytica ar fi folosit alte strategii decât fictiva mea agenţie de voiaje de lux, Fry’s. Abordarea lor a constat în identificarea unor grupuri reduse de persoane pe care le considerau uşor de convins şi cărora le trimiteau apoi reclame ţintite, vizându-le direct, în loc să le trimită reclame nediferenţiate, pe baza unui contract-cadru[26]. S-a descoperit, de exemplu, că anumite categorii se suprapuneau într-o mare măsură, spre exemplu, aceea a persoanelor care cumpăraseră produse americane get-beget precum autoturisme Ford şi aceea a cetăţenilor care se înregistraseră în registrele de votare ca susţinători ai Partidului Republican. Aşadar, cei de la Cambridge Analytica au început să caute persoane care preferau marca Ford, dar despre care nu se ştia şi că ar fi votat cu republicanii, să vadă dacă puteau să le influenţeze opiniile prin utilizarea de reclame naţionaliste, care să le stârnească patriotismul. Într-un fel, asta nu diferă prea mult de strategia folosită de un candidat care identifică un cartier cu alegători indecişi şi se duce din uşă în uşă pentru a-i convinge pe rând. Iar în materie de

26 Blanket advertising (engl.), un aranjament de cumpărare de reclame, în care o agenţie de publicitate se obligă să distribuie un anumit număr, la un anumit preţ fix, pe o perioadă determinată, (n. tr.).

campanii on-line, nu se deosebeşte de ceea ce au făcut Obama şi Clinton în timpul propriilor campanii electorale. Fiecare partid politic important din lumea occidentală foloseşte analize extensive ale datelor informatice şi diverse modalităţi de micro-targetare a alegătorilor.

Dacă este să dăm crezare reportajului sub acoperire înregistrat de *Channel Four News*, atunci Cambridge Analytica a folosit şi strategia realizării de profiluri de personalitate ale electoratului pentru difuzarea de mesaje politice cu mare încărcătură emoţională – de exemplu, au descoperit mame singure care aveau într-un procent mai mare un caracter nevrotic şi au profitat de teama lor de a fi atacate în propria locuinţă, pentru a le convinge să susţină un mesaj politic care făcea lobby pentru deţinerea armelor de foc. Agenţiile de publicitate au folosit în mod cert pe scară largă aceste stratageme şi probabil că la fel au făcut-o şi coordonatorii altor campanii politice.

Numai că pe lângă asta, Cambridge Analytica este acuzată şi de crearea de reclame şi mascarea lor în jurnalism. Conform declaraţiilor unui denunţător, preluate de ziarul britanic *Guardian*, una dintre cele mai eficiente reclame din timpul campaniei a fost un grafic interactiv intitulat: „10 adevăruri incomode despre Fundaţia Clinton”.21 Un alt denunţător a mers şi mai departe, susţinând că „articolele”.

plantate de Cambridge Analytica în presă se bazau, adesea, pe minciuni demonstrabile.22

Să presupunem, de dragul discuţiei, că toate cele de mai sus ar fi adevărate: Cambridge Analytica a transmis diverse ştiri false (*fake news*) manipulatoare pe Facebook anumitor persoane, în funcţie de profilurile psihologice ale acestora, întrebarea este: a dat roade strategia lor?

MICRO-MANIPULAREA

Există o totală lipsă de simetrie în modul în care ne raportăm la puterea pe care o deţin reclamele politice ţintite. Ne place să credem că noi înşine avem o gândire independentă, că suntem imuni la manipulare şi, totuşi, ni-

i imaginăm pe „ceilalţi” – în special, pe cei care au convingeri politice diferite – ca fiind inimaginabil de creduli. Adevărul este, probabil, undeva, la mijloc.

Ceea ce ştim în mod sigur este că postările pe care le vedem pe Facebook au capacitatea de a ne modifica emoţiile. În cadrul unui experiment controversat, realizat de angajaţi ai Facebook în anul 2013, a manipulat fluxurile de ştiri care li se afişau unui număr de 689 003 utilizatori fără cunoştinţa (sau consimţământul) acestora, în încercarea de a le controla emoţiile şi de a le influenţa starea de spirit.23 Cei care au realizat cercetarea au blocat toate postările prietenilor respectivilor utilizatori care conţineau cuvinte cu semnificaţii pozitive, după care au făcut acelaşi lucru cu postările care includeau cuvinte negative şi au urmărit cum reacţionau subiecţii neavizaţi în fiecare dintre aceste situaţii. Utilizatorii care au observat mai puţin conţinut negativ în fluxurile lor au început să posteze ei înşişi lucruri mai degrabă pozitive. Între timp, cei cărora le fuseseră ascunse din *timeline* postările pozitive au început să folosească ei înşişi cuvinte negative. Concluzia: ne putem imagina c-am fi imuni la manipularea emoţională, dar probabil că nu suntem.

Mai ştim, din experimentul lui Epstein descris în capitolul „Puterea”, că până şi ordinea paginilor care afişează rezultatele unui motor de căutare poate fi suficientă pentru a-i influenţa pe alegătorii indecişi să devină favorabili unui anumit candidat şi nu altuia. Cunoaştem, de asemenea, din activitatea aceloraşi cercetători universitari ai căror algoritmi Cambridge Analytica i-a utilizat într-un scop diferit, că reclamele sunt mai eficiente atunci când vizează trăsăturile de personalitate ale indivizilor.

Toate aceste informaţii însumate constituie un argument puternic pentru a sugera că astfel de metode pot influenţa modul în care oamenii votează, tot aşa cum au o influenţă asupra modului în care îşi cheltuiesc banii. Dar – şi este un mare „dar” –, mai trebuie să ştiţi ceva înainte de

a trage o concluzie.

Toate cele de mai sus sunt adevărate, numai că efectele propriu-zise sunt nesemnificative. În experimentul realizat de cei de la Facebook, utilizatorii au fost, într-adevăr, mai înclinaţi să posteze mesaje pozitive dacă li se ascundeau ştirile cu conţinut negativ. Numai că diferenţa era de mai puţin de o zecime din 1%!

În mod asemănător, în exemplul cu reclamele personalizate, produsele cosmetice vândute introverţilor aveau mai mult succes dacă se luau în considerare trăsăturile de personalitate ale unei persoane, însă diferenţa obţinută era infimă. O reclamă generică determina 31 de persoane dintr-o mie să dea clic pe ea. Chiar şi procentul acela de 50% îmbunătăţire pe care l-am citat mai devreme, care era îndrăzneţ întipărit la începutul lucrării ştiinţifice, se referă, de fapt, la o creştere infimă, de la 11 clicuri la 1.000 de persoane care văzuseră reclama, la 16!

Strategiile pot să dea rezultate, este adevărat. Dar companiile de publicitate nu-şi injectează mesajele direct în minţile unei audienţe pasive. Nu suntem nişte ţinte chiar atât de uşoare. Ne pricepem mult mai bine să ignorăm reclamele sau să interpretăm în felul nostru mesajele propagandiste decât şi-ar dori cei care ni le trimit. Până la urmă, chiar şi în cazul celor mai bune şi mai parşive campanii publicitare axate pe micro-profilare, grupul-ţintă va ajunge să fie influenţat într-o măsură foarte mică.

Cu toate acestea, la nivel teoretic, cel puţin, într-o campanie electorală, astfel de frânturi de capacitate de influenţare ar putea fi suficiente pentru a înclina balanţa în favoarea unui candidat sau a altuia. La o populaţie de zeci sau sute de milioane de indivizi, acele schimbări de opinie de unu la mie se pot însuma cu repeziciune. Şi, când îţi aminteşti că, aşa cum a subliniat Jamie Bartlett într-un articol pentru revista britanică *Spectator*, Trump a câştigat alegerile din statul Pennsylvania cu o diferenţă de 44.000 de voturi „pentru” din 6.000.000 exprimate, pe cele din

Wisconsin, cu 22.000 de voturi, iar pe cele din Michigan cu doar 11.000 de voturi, atunci s-ar putea ca până şi nişte marje de influenţare de sub 1% să fie exact ceea ce-ţi trebuie.24

Adevărul este că nu se poate preciza ce efect au avut toate acestea în alegerile prezidenţiale din Statele Unite. Chiar dacă am avea acces la toate faptele, nu ne putem întoarce în timp, să descâlcim păienjenişul de cauze şi efecte pentru a indica fie şi un singur motiv pentru decizia de a vota a vreunei persoane. Ce-a fost a fost şi nu mai poate fi schimbat. Ce contează este spre ce ne îndreptăm în viitor.

EVALUEAZĂ-MĂ!

Este important să ne amintim că am beneficiat cu toţii de acest model pe care este construit internetul. Peste tot în lume, oamenii au acces liber şi gratuit la reţelele globale de comunicare instantanee, având tot tezaurul cunoaşterii umane la îndemână, informaţii de ultimă oră din toate colţurile pământului şi utilizare nelimitată a celor mai remarcabile programe de calculator şi a celor mai avansate tehnologii, realizate de companii private – totul plătit de publicitate. Aceasta este înţelegerea pe care am încheiat-o. Tehnologie gratuită, în schimbul posibilităţii ca aceasta să fie folosită pentru a ne influenţa şi pentru a se profita de pe urma noastră. Tot ce este mai bun şi tot ce este mai rău din capitalism dintr-o singură mişcare simplă.

Am putea să ajungem la concluzia că suntem fericiţi cu târgul pe care l-am încheiat. Şi asta este cât se poate de în regulă. Dacă îl acceptăm, însă, este important să fim conştienţi de la bun început de toate pericolele pe care le poate genera strângerea atâtor informaţii. Trebuie să avem în minte la ce ar putea duce aceste baze de date în viitorchiar dincolo de problemele legate de protecţia datelor şi de potenţialul lor de a submina democraţia (de parcă acestea nu ar fi, oricum, îndeajuns de rele!). Dar există şi o altă răsturnare de situaţie în această poveste distopică: o aplicaţie creată pentru a folosi astfel de baze

de date detaliate, uriaşe şi intercorelate, care pare demnă să fie prezentată într-un episod al popularului serial SF de pe Netflix *Black Mirror*[27], dar care există în realitate. Este cunoscută drept Sesame Credit, este utilizată de guvernul chinez şi reprezintă un sistem care acordă un punctaj fiecărui cetăţean.

Imaginaţi-vă că fiecare informaţie pe care un broker de date ar avea-o despre dumneavoastră ar fi redusă la un punctaj unic. Absolut tot a fi inclus: cât de responsabil (sau de iresponsabil) v-aţi achitat creditele, numărul telefonului mobil, adresa – lucrurile obişnuite. Dar şi comportamentul de zi cu zi, postările de pe reţelele de socializare, informaţiile adunate de aplicaţia prin care solicitaţi taxiuri, chiar şi înregistrări ale activităţii on-line pe un site de anunţuri matrimoniale. Rezultatul este un singur număr cuprins între 350 şi 950 de puncte.

Sesame Credit nu dezvăluie detalii despre algoritmul său „complex" de punctare. Însă Li Yingyun, directorul serviciului de tehnologie, a oferit câteva exemple despre ceea ce s-ar putea deduce din rezultatele obţinute, într-un interviu luat de *Caixin Media*, grupul mediatic cu sediul în Beijing:

Cineva care joacă jocuri video zece ore pe zi, de exemplu, va fi considerat o persoană inactivă. Cineva care cumpără frecvent scutece pentru copii va putea fi considerat părinte, care, prin comparaţie, este mai probabil să posede un anumit simţ al responsabilităţii.25

Iar pentru chinezi, acest scor chiar a ajuns să conteze. Dacă punctajul tău depăşeşte 600 de puncte, poţi să obţii un card de credit special. Dacă depăşeşte 666 de puncte, eşti recompensat cu o limită de credit superioară. Cei cu punctaje peste 650 pot să închirieze o maşină fără

27 „Nosedive", primul episod al celui de-al treilea sezon al serialului Black Mirror (în regia lui Joe Wright şi scenariul lui Michael Schur şi Rashida Jones, după o poveste de Charlie Brooker), a fost adesea considerat o metaforă pentru Social Credit System, sistemul de credit social implementat de guvernul chinez, preuspoziţie confirmată de producătorul Charlie Brooker într-o conferinţă de presă (n. red.).

să depună garanţie şi să folosească un culoar pentru VIP-uri la aeroportul din Beijing. Oricine are un punctaj de peste 750 poate să solicite în regim de urgenţă o viză de intrare în Europa.26

Totul este distractiv acum, cât timp sistemul de acordare a unor punctaje cetăţenilor este încă opţional. Dar în 2020, când va deveni obligatoriu, persoanele cu punctaje mici vor începe să simtă repercusiunile în fiecare aspect al vieţii lor. Chiar documentele guvernului cu privire la acest sistem prezintă exemple de sancţiuni, care s-ar putea aplica oricărei persoane considerate recalcitrante:

Restricţii în privinţa plecării în străinătate, restricţii de cumpărare a unor […] bunuri, de a călători cu avionul, în scop de turism şi în vacanţe, sau restricţii de cazare în hoteluri cu un anumit număr de stele.

Documentele oficiale mai avertizează că, în cazul unor „subiecţi care vor înşela grav încrederea [partidului]”, acesta va „sfătui băncile comerciale […] să le limiteze respectivilor accesul la credite, la cumpărarea de asigurări şi la alte servicii financiare de acelaşi gen”.27 Loialitatea este elogiată. Înşelarea încrederii este pedepsită. Aşa cum se exprima Rogier Creemers, un cadru universitar specializat în dreptul şi guvernarea chineză care predă la Institutul „Van Vollenhoven” al Universităţii Leiden:

Modalitatea cea mai bună de a înţelege acest sistem este să-l considerăm un soi de bastard, rod al iubirii pătimaşe, al unui sistem de fidelizare a clienţilor.28

Nu pot să vă ofer informaţii prea liniştitoare în privinţa lui Sesame Credit, dar nu doresc nici să credeţi că perspectivele noastre sunt cu totul şi cu totul sumbre, căci se întrezăreşte o rază de speranţă în altă parte. Oricât de mohorât ar părea viitorul, sunt semnale că, încetul cu încetul, direcţia vântului se schimbă. Mulţi dintre cei care fac parte din comunitatea specialiştilor în ştiinţa datelor sunt, de ceva timp, la curent cu şi obiectează la

exploatarea informaţiilor despre oameni în schimbul obţinerii de profituri. Până la scandalul stârnit de Cambridge Analytica, însă, aceste chestiuni nu au atras atenţia în mod special, neajungând, nicăieri în lume, pe prima pagină a ziarelor. La începutul lui 2018 însă, când a izbucnit scandalul, publicul larg a văzut pentru întâia oară cum algoritmii culeg, pe tăcute, informaţii despre oameni şi şi-a dat seama că, în lipsa unei supravegheri atente sau a unor reglementări, acest lucru ar putea avea repercusiuni dramatice.

Iar reglementările nu întârzie să sosească. Dacă locuiţi în Uniunea Europeană, aici a fost adoptată de curând o nouă lege, numită GDPR (General Data Protection Regulation[28]), care decretează a fi ilegale destul de multe dintre activităţile întreprinse de către brokerii de date. Teoretic, acestora nu li se va mai permite să vă stocheze informaţiile fără un scop explicit. Nu vor putea să deducă informaţii despre persoana dumneavoastră în lipsa consimţământului exprimat explicit. Şi nu vor mai putea să vă ceară permisiunea de a vă colecta datele personale dintr-un anumit motiv, iar apoi să le folosească în secret pentru altceva. Asta nu înseamnă neapărat că acest gen de practici vor şi înceta. În primul rând, pentru că, de cele mai multe ori, noi nici nu dăm atenţie termenilor şi condiţiilor pentru care semnăm că suntem de acord atunci când facem clic on-line în stânga şi în dreapta, aşa că ne-am putea găsi în situaţia de a ne exprima consimţământul chiar fără să ne dăm seama. În al doilea rând, identificarea practicilor ilegale şi aplicarea reglementărilor rămâne o chestiune complicată într-o lume în care cea mai mare parte a analizei şi transferului de date se face din umbră. Prin urmare, trebuie să aşteptăm şi să vedem cum vor decurge lucrurile.

28 Regulamentul general privind protecţia datelor (denumire completă: Regulamentul nr. 679 din 27 aprilie 2016 privind protecţia persoanelor fizice în ceea ce priveşte prelucrarea datelor cu caracter personal şi libera circulaţie a acestor informaţii şi de abrogare a Directivei 95/46/CE) (n. tr.).

Europenii sunt mai norocoşi, dar şi în Statele Unite există persoane care fac presiuni pentru reglementarea situaţiei, în 2014, Federal Trade Comission (Comisia Federală pentru Comerţ) a publicat un raport în care condamna practicile dubioase ale brokerilor de date, iar de atunci a exercitat tot felul de presiuni pentru a spori drepturile consumatorilor. Apple a integrat deja un soft pentru „protecţie inteligentă împotriva urmăririi”[29] în browserul său, Safari, iar Firefox a făcut acelaşi lucru. Facebook îşi întrerupe legăturile cu brokerii săi de informaţii. Argentina şi Brazilia, Coreea de Sud şi multe alte ţări au impus legi asemănătoare GDPR-ului. Poate că Europa este cu un pas înainte, dar se manifestă o tendinţă globală de a încerca să ne îndreptăm în direcţia cea bună.

Dacă datele personale informatizate sunt, cu adevărat, noul aur înseamnă că, până acum, am trăit în Vestul Sălbatic.

Eu, una, sunt, însă, optimistă şi cred că, pentru cei mai mulţi dintre noi, ce-a fost mai rău a trecut deja.

Am face, totuşi, bine să ne amintim că în viaţă nimic nu este gratis. În timp ce legislaţia încearcă să recupereze decalajul faţă de realitatea tehnologică şi se poartă, ca întotdeauna, bătălia cu dorinţa de profit a corporaţiilor pentru obţinerea a ceea ce este bine pentru societate, trebuie să fim atenţi să nu ne amăgim cu falsa impresie că am avea parte de confidenţialitate. Ori de câte ori folosim un algoritm – mai ales unul care ne este oferit fără să plătim! – trebuie să ne întrebăm care sunt motivaţiile ascunse ce stau la baza acestei acţiuni. De ce această aplicaţie îmi oferă astfel de chestii în mod gratuit? Ce face, de fapt, algoritmul acesta? Este, oare, un schimb care nu mă deranjează? Mi-ar fi mai bine fără el?

Este o lecţie valabilă şi dincolo de mediul virtual, pentru, că, în prezent, influenţa acestui gen de calcule se extinde, practic, asupra oricărui aspect al societăţii. Informaţiile şi algoritmii nu mai deţin doar puterea de a

29 Intelligent Tracking Prevention sau ITP (n. tr.).

prezice obiceiurile la shopping ale cuiva. Ele pot, de asemenea, să priveze o persoană de liberate.

Capitolul 3 JUSTIŢIA

Vara, nu este deloc neobişnuit să dai peste petrecăreţi inofensivi care beau câte un pahar, într-o seară de duminică, pe străzile din Brixton, acolo unde începe următoarea noastră poveste. Acesta este un district din sudul Londrei, care are reputaţia de a fi un loc bun în care să te distrezi cu prietenii. Iar într-una din seri tocmai se încheiase un festival de muzică, iar zona era plină de oameni care se pregăteau, bine dispuşi, să se întoarcă acasă sau să continue cheful. Pe la 23.30, însă, starea lor de spirit s-a schimbat. În curtea unei locuinţe sociale a izbucnit o bătaie, iar când poliţiştii nu au reuşit să ţină situaţia sub control, totul s-a extins rapid până în centrul localităţii, unde sute de tineri li s-au alăturat recalcitranţilor.

Era în august 2011. Cu o noapte înainte, în cealaltă parte a oraşului, un protest care începuse paşnic împotriva împuşcării de către poliţie a unui tânăr din Tottenham pe nume Mark Duggan se transformase într-o confruntare violentă. Acum, pentru a doua noapte la rând, zone întregi din oraş erau cuprinse de haos, iar de data aceasta atmosfera era diferită. Ceea ce începuse ca o demonstraţie locală se transformase repede într-o amplă manifestare a sfidării legii şi ordinii şi o devalizare generalizată.

Chiar când răzmeriţa începea să ia amploare, Nicholas Robinson, un student la Electrotehnică în vârstă de 23 de ani ce îşi petrecuse sfârşitul de săptămână la prietena lui, se întorcea acasă folosind scurtătura sa obişnuită prin Brixton.1 De acum, străzile familiare deveniseră, practic, de nerecunoscut: de o parte şi de alta a drumului se zăreau maşini răsturnate, ferestre sparte, case incendiate şi magazine vandalizate.2 Poliţiştii se străduiau, disperaţi, să calmeze situaţia, dar le era cu neputinţă să oprească maşinile şi scuterele care veneau chiar lângă vitrinele sparte şi încărcau bunurile de furat:

haine, încălţăminte şi televizoare. Brixtonul scăpase complet de sub control.

La câteva străzi distanţă de un magazin de produse electronice care fusese complet devastat, Nicholas Robinson a trecut pe lângă supermarketul unde îşi făcea de obicei cumpărăturile. Aproape la fel ca şi celelalte magazine, şi acesta era un dezastru: vitrinele şi uşile din sticlă fuseseră sparte, iar rafturile din interior erau vraişte din cauza distrugerilor provocate de jefuitori. Valuri de insurgenţi treceau în fugă pe lângă el, ţinând strâns în braţe laptopuri noi-nouţe, fără a fi opriţi de poliţie. În mijlocul haosului, simţind că-i este sete, Nicholas a intrat în magazin şi a luat un bax cu sticle de apă în valoare de 3, 50 lire sterline. Şi, tocmai când dădea colţul, pregătindu-se să plece, poliţiştii au dat buzna în supermarket. Nicholas şi-a dat imediat seama ce făcuse, a lăsat să-i cadă baxul cu sticle şi a încercat s-o zbughească de-acolo.3

Când a sosit noaptea de luni, ţara se pregătea să înfrunte noi răzmeriţe. Şi, într-adevăr, în noaptea aceea, jefuitorii au ieşit din nou pe străzi.4 Printre aceştia se număra şi tânărul de 19 ani Richard Johnson. Intrigat de ceea ce văzuse la ştiri, pusese mâna pe o cagulă (un articol cu totul nepotrivit pentru sezonul de vară!), sărise în maşină şi-şi croise drum până la centrul comercial din apropiere. Cu faţa ascunsă, Richard a pătruns în fugă în magazinul de jocuri pe computer al oraşului, a adunat un braţ de jocuri pe calculator şi s-a întors la maşină.5 Din păcate pentru el, parcase chiar în faţa unei camere CCTV[30]. Numărul de pe plăcuţa de înmatriculare a maşinii a ajutat poliţia să-l urmărească, iar înregistrarea video a făcut mai uşoară deschiderea unui proces împotriva lui.6

Atât Richard Johnson, cât şi Nicholas Robinson au fost arestaţi pentru faptele comise de ei în timpul revoltelor din 2011. Ambii au fost acuzaţi de tâlhărie. Ambii s-au prezentat în faţa judecătorilor. Ambii s-au

30 Closed-circuit television (engl.), „camere de televiziune cu circuit închis”, folosite pentru supravegherea video a unei zone (n. tr.).

declarat vinovaţi. Însă aici iau sfârşit asemănările dintre cele două procese.

Nicholas Robinson a fost chemat cel dintâi pe banca acuzaţilor, apărând în faţa unui judecător de la Curtea Magistraţilor din Camberwell la mai puţin de o săptămână de la incident. În pofida valorii mici a sticlelor de apă pe care le furase, în pofida absenţei oricăror antecedente penale, a faptului că urma cursurile învăţământului la zi şi că a declarat curţii că îi era ruşine de faptele pe care le comisese, judecătorul a decis că acţiunile sale contribuiseră la atmosfera de anarhie din Brixton din acea noapte. Şi, astfel, în plânsetele membrilor familiei sale din spaţiul sălii de judecată rezervat publicului, Nicholas Robinson a fost condamnat la şase luni de închisoare.7

Cazul lui Johnson a ajuns în instanţă în ianuarie 2012. Cu toate că ieşise din casă purtând un obiect de îmbrăcăminte menit să-i ascundă identitatea din dorinţa deliberată de a jefui, cu toate că şi el jucase un rol în agravarea dezordinii publice, Johnson a scăpat de închisoare. I s-a dat o sentinţă cu suspendare şi a fost condamnat la efectuarea a 200 de ore de muncă neplătită în folosul comunităţii.8

DILEMA MORALĂ A CONSECVENŢEI

Sistemul judiciar ştie că nu este perfect, însă nici nu încearcă să fie astfel. Aprecierea vinovăţiei şi stabilirea pedepselor nu este o ştiinţă exactă şi este imposibil ca un judecător să garanteze că sentinţele sale vor fi date cu precizie. De aceea şi există expresii precum „suspiciune rezonabilă” şi „motive temeinice”, atât de vitale pentru vocabularul legislativ, şi tot din aceste motive recursurile sunt o parte atât de importantă a proceselor: sistemul recunoaşte astfel că este imposibil de atins certitudinea absolută asupra vinovăţiei sau nevinovăţiei cuiva.

Chiar şi aşa, discrepanţele în tratamentul aplicat unor inculpaţi – precum Nicholas Robinson şi Richard Johnson – par incorecte. Sunt prea mulţi factori implicaţi ca să se poată afirma cu siguranţă că existenţa unei

diferenţe între sentinţe este „nedreaptă”, dar, în limite rezonabile, s-ar putea spera că judecătorii au fost, în general, consecvenţi în modul în care au luat deciziile. Dacă, de exemplu, dumneavoastră şi geamănul dumneavoastră aţi comis aceeaşi infracţiune, aţi putea spera că o instanţă de judecată vă va da aceeaşi sentinţă. Însă, aşa să fie, oare?

În anii 1970, un grup de cercetători americani a încercat să răspundă la o variantă a acestei întrebări.9 în loc să folosească infractori gemeni (ceea ce este, practic, dificil şi, din punct de vedere moral, indezirabil), oamenii de ştiinţă au inventat o serie de cazuri ipotetice şi au întrebat, separat, 47 de judecători din cadrul instanţelor districtuale ale statului Virginia cum ar proceda în fiecare dintre ele. Iată un exemplu extras din studiu, ca să încercaţi acelaşi lucru. Cum aţi proceda în cazul următor?

O inculpată, o tânără de 18 ani, a fost reţinută pentru posesie de marijuana, fiind arestată împreună cu prietenul ei şi cu alte şapte cunoştinţe. Au fost găsite cantităţi însemnate de marijuana, filmată şi nefiimată, însă asupra fetei nu s-a descoperit nici urmă.

Nu avea antecedente penale, era o studentă silitoare provenind dintr-o familie din clasa mijlocie şi nu fusese nici sfidătoare, dar nici nu-şi ceruse scuze pentru faptele ei.

Deosebirile între sentinţe au fost impresionante. Din cei 47 de judecători, 29 au declarat inculpata nevinovată, iar 18 – vinovată. Dintre judecătorii care au optat pentru verdictul „vinovată”, opt au recomandat eliberarea condiţionată, patru s-au gândit că o amendă era mai potrivită, trei ar fi optat şi pentru eliberare condiţionată, şi pentru amendă, iar trei ar fi fost înclinaţi s-o trimită pe inculpată la închisoare.

Aşadar, în baza aceloraşi dovezi şi în cazuri identice, un inculpat se poate aştepta să scape basma curată sau să fie trimis direct la închisoare, asta depinzând cu totul de judecătorul în faţa căruia a avut şansa (sau neşansa) să se prezinte.

Este o adevărată lovitură pentru oricine se mai agaţă încă de ideea unei justiţii consecvente. Dar lucrurile stau chiar mai rău, pentru că judecătorii nu numai că sunt în dezacord unul cu altul, ci mai au şi tendinţa să-şi contrazică până şi *propriile sentinţe*.

Într-un studiu recent, 81 de judecători din Marea Britanie au fost întrebaţi dacă ar oferi eliberarea pe cauţiune unui număr de inculpaţi imaginari.10 Fiecare caz ipotetic avea propria poveste fictivă şi antecedente penale inventate. Asemenea confraţilor lor din studiul realizat în statul Virginia, judecătorii britanici nu au reuşit să cadă de acord nici măcar asupra unui singur caz din cele 41 care le-au fost prezentate.11 De data aceasta, printre cele 41 de cazuri ipotetice prezentate fiecărui judecător, şapte dintre ele apăreau de două ori – cu numele inculpaţilor schimbate la a doua apariţie pentru ca judecătorul să nu-şi dea seama că erau duplicate. Era un şiretlic parşiv, însă revelator. Cei mai mulţi judecători nu au reuşit să ia aceeaşi decizie în acelaşi dosar atunci când l-au văzut pentru a doua oară. Uimitor este că unii judecători nu s-au descurcat mai bine în a da acelaşi răspuns într-o situaţie identică decât dacă ar fi stabilit, efectiv, cauţiunea la întâmplare.12

Multe alte studii au ajuns la o concluzie identică: ori de câte ori judecătorii au libertatea de a evalua cazurile ei înşişi, vor apărea neconcordanţe majore. A lăsa judecătorilor libertate de interpretare înseamnă a permite existenţa elementului şansă în sistem.

Există o soluţie simplă, desigur. O cale uşoară pentru a garanta consecvenţa judecătorilor este aceea de a le lua dreptul de a exercita puteri discreţionare. Dacă fiecare persoană acuzată de acelaşi delict ar fi tratată în exact aceeaşi manieră, atunci s-ar putea garanta acurateţea actului de justiţie – cel puţin, în ceea ce priveşte stabilirea cauţiunii şi a pedepsei. Şi, într-adevăr, unele ţări au pornit pe calea aceasta. În Statele Unite, la nivel federal, şi în unele zone din Australia operează deja sisteme prescriptive de stabilirea sentinţelor.13 însă acest gen de

consecvenţă are, totuşi, neajunsuri, căci, pentru tot ce se câştigă în precizie, se pierde în ceea ce priveşte un alt fel de echitate.

Pentru a ilustra ceea ce am spus, închipuiţi-vă doi inculpaţi, ambii acuzaţi de furt dintr-un supermarket. Unul este un infractor de carieră, oarecum în elementul lui, care fură de bunăvoie; celălalt şi-a pierdut de curând locul de muncă şi se luptă să o scoată la capăt, furând ca să-şi hrănească familia şi fiind plin de remuşcări din cauza faptelor sale. Prin eliminarea libertăţii de a ţine cont de aceste circumstanţe atenuante, regulile stricte care îi tratează în acelaşi fel pe toţi cei care se fac vinovaţi de aceleaşi infracţiuni pot sfârşi prin a fi inutil de aspre pentru unii şi pot însemna că se ratează orice şansă de a-i reabilita pe anumiţi infractori.

Iată o adevărată dilemă morală. Oricum s-ar stabili sistemul în care să acţioneze judecătorii, trebuie găsit un echilibru delicat între oferirea unei justiţii individualizate şi asigurarea consecvenţei sentinţelor. Cele mai multe ţări au încercat să rezolve dilema optând pentru adoptarea unor sisteme care se situează undeva între extrema prescriptivă a legii federale din Statele Unite şi aceea în care totul se bizuie în întregime pe puterea discreţionară a judecătorilor – cum este cazul în Scoţia.14 în întreaga lume occidentală, regulile privind pronunţarea sentinţelor tind să stabilească o sentinţă maximă (ca în Irlanda), o sentinţă minimă (ca în Canada) sau ambele (ca în Anglia şi Ţara Galilor) 15 şi lasă la latitudinea judecătorilor ajustarea sentinţei în plus sau în minus între aceste limite.

Niciun sistem nu este perfect. Există întotdeauna un amestec de nedreptăţi care se iau la întrecere, un haos al injustiţiilor care se opun una alteia. Dar în mijlocul acestui conflict şi cu toată această complexitate, un algoritm are prilejul să fie în elementul său. Pentru că – în mod remarcabil – cu un algoritm care ar face parte din proces se pot garanta atât consecvenţa, cât şi justiţia în fiecare caz în parte. Nimeni nu mai este nevoit să aleagă între cele

două.

ECUAŢIA JUSTIŢIEI

Algoritmii nu pot lua decizii în ceea ce priveşte vinovăţia cuiva. Ei nu pot să cântărească argumentele aduse de apărare şi de acuzare, nici să studieze dovezile şi nici să stabilească dacă un inculpat are într-adevăr remuşcări. Prin urmare, nu trebuie să ne aşteptăm ca aceştia să înlocuiască judecătorii în viitorul apropiat. Ceea ce poate să facă un algoritm, oricât de incredibil ar părea, este să proceseze informaţiile despre o anumită persoană şi să calculeze care este riscul ca aceasta să recidiveze. Iar, din moment ce deciziile judecătorilor se bazează pe probabilitatea ca un inculpat să comită din nou o infracţiune, aceasta se dovedeşte o aptitudine destul de utilă.

Informaţiile şi algoritmii se folosesc în sistemul judiciar de aproape un secol, primele exemple datând din America anilor 1920. Pe atunci, în sistemul american sentinţele infractorilor condamnaţi aveau o durată maximă, iar, după o anumită perioadă, aceştia deveneau eligibili pentru eliberarea condiţionată[31]. Zeci de mii de prizonieri au beneficiat astfel de eliberarea înainte de termen. Unii au fost reabilitaţi cu succes, alţii nu. Însă, împreună, au constituit „baza de date” perfectă pentru organizarea unui experiment menit să descopere: este posibil să prezici dacă un deţinut îşi va respecta sau nu condiţiile eliberării condiţionate?

Şi acum îşi face intrarea în scenă Ernest W. Burgess, un sociolog canadian de la Universitatea din Chicago pasionat de predicţii. Burgess a fost un mare susţinător al cuantificării fenomenelor sociale. În cursul carierei sale, a încercat să prognozeze totul – de la efectele pensionării şi până la reuşita în căsnicie iar în 1928 a devenit prima

31 Parole (engl.). Un fapt amuzant: „parole” vine din francezul parole, care înseamnă „cuvinte vorbite”. Se regăseşte în forma originală prin anii 1700, când prizonierii erau eliberaţi dacă îşi dădeau cuvântul de onoare că nu vor recidiva. Sursa: https://www.etymonline.com/word/parole (n. a.).

persoană care a conceput cu succes un instrument pentru calcularea riscului de apariţie a comportamentului infracţional, unul care se baza mai mult pe măsurători decât pe intuiţie.

Folosind diverse informaţii despre 3.000 de deţinuţi din trei închisori din statul Illinois, Burgess a identificat 21 de factori pe care i-a considerat „posibil semnificativi" pentru a determina şansele ca un deţinut să îşi încalce termenii eliberării condiţionate. Printre aceştia se numărau: genul infracţiunii, numărul de luni petrecute în închisoare şi tipul social al deţinutului, pe care – cu toată „delicateţea" la care ne putem aştepta din partea unui sociolog de la începutul secolului al XX-lea – i-a împărţit în categorii precum „vagabond", „beţiv", „om de nimic", „băiat de la fermă" şi „imigrant".16

Burgess i-a alocat fiecărui deţinut un punctaj între zero şi unu la fiecare dintre aceşti douăzeci şi unu de factori. Pe bărbaţii care obţineau punctaje mari (între 16 şi 21) îi considera a fi mai puţin predispuşi să recidiveze, pe aceia care aveau punctaje mici (patru sau mai puţin) îi considera susceptibili să îşi încalce condiţiile de eliberare condiţionată.

Când toţi condamnaţii au fost eliberaţi, fiind, prin urmare, liberi să încalce condiţiile de eliberare condiţionată dacă ar fi ales să facă acest lucru, Burgess a avut ocazia să verifice cât de precise fuseseră predicţiile sale. Cu ajutorul unei analize atât de simpliste, el a reuşit să fie remarcabil de exact. 98% dintre aceia aflaţi în grupul său cu risc scăzut au reuşit să depăşească cu bine perioada de eliberare condiţionată, iar două treimi dintre cei din grupul considerat ca având un risc ridicat au eşuat.17 După cum s-a dovedit, până şi modelele statistice primitive au putut face predicţii mai precise decât specialiştii.

Însă activitatea lui Burgess şi-a avut criticii săi. Observatori sceptici s-au întrebat în ce măsură factorii care preziceau în mod fiabil succesul respectării condiţiilor

de eliberare condiţionată într-un anumit loc şi într-un anumit moment puteau fi aplicaţi în altă parte. (Aveau dreptate: nu sunt foarte convinsă că, de exemplu, categoria „băiat de la fermă” ar fi de mare ajutor în estimarea pericolului de recidivare printre infractorii urbani din zilele noastre.) Alţi oameni de ştiinţă l-au criticat pe Burgess pentru că nu făcea decât să utilizeze toate informaţiile de care dispunea, fără să cerceteze dacă acestea erau sau nu relevante.18 S-au mai pus şi întrebări referitoare la modul în care acorda el punctaje deţinuţilor: în definitiv, metoda sa era ceva mai mult decât o opinie tradusă în ecuaţii. Nu este mai puţin adevărat că puterea lui de previziune a fost atât de impresionantă, încât, în 1935, metoda Burgess a pătruns în închisorile din statul Illinois pentru a ajuta comisiile de eliberare condiţionată să ia decizii.19 Iar pe la începutul secolului al XXI-lea, peste tot în lume se utilizau modele matematice care erau „urmaşele” metodei lui Burgess.20

Revenind la epoca modernă, algoritmii de ultimă generaţie folosiţi pentru evaluarea riscurilor în curţile de judecată sunt mult mai sofisticaţi decât instrumentale rudimentare inventate de Burgess. Aceştia nu vin numai în sprijinul deciziilor de eliberare condiţionată, ci sunt întrebuinţaţi şi ca să ajute la repartizarea corectă a deţinuţilor în diverse programe de intervenţie, pentru a decide cui să i se ceară cauţiune şi, de puţin timp, pentru a sprijini judecătorii în luarea deciziilor de condamnare. Principiul lor fundamental de funcţionare este cel dintotdeauna: în sistem intră informaţiile despre acuzat – vârstă, antecedente penale, gravitatea infracţiunii comise şi aşa mai departe – şi iese o predicţie referitoare la cât de mare ar fi riscul de a-l lăsa în libertate.

Aşadar, cum funcţionează totul? Ei bine, în mod general, algoritmii cei mai performanţi din prezent folosesc o tehnică numită *random forest* (» pădure aleatorie”), care, în esenţă, are la bază o idee extrem de

simplă: umilul arbore de decizie.[32]

ÎNTREABĂ PUBLICUL

S-ar putea să fiţi familiarizaţi cu arborii de decizie din ceea ce aţi învăţat la şcoală. Ca metodă de structurare a observaţiilor, precum darea cu banul sau aruncarea zarurilor, aceştia sunt populari în rândul profesorilor de matematică. Odată construit, un arbore de decizie poate fi folosit ca o diagramă: se ia un set de împrejurări şi se analizează pas cu pas ce trebuie făcut sau, în cazul de faţă, ce se va întâmpla.

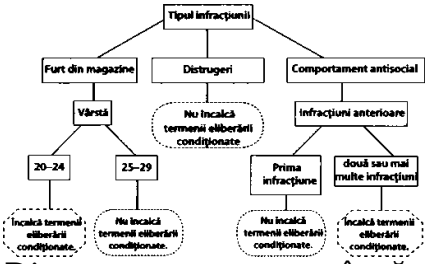
Imaginaţi-vă că încercaţi să hotărâţi dacă să cereţi cauţiune pentru o anumită persoană. Ca şi în cazul eliberării condiţionate, această decizie se bazează pe un calcul clar. Vinovăţia este irelevantă. Nu trebuie decât să faceţi o predicţie: va încălca inculpatul condiţiile eliberării pe cauţiune dacă i se va permite să iasă din închisoare?

Ca să vă ajute să faceţi o predicţie, dispuneţi de informaţiile culese de la un număr redus de recidivişti, dintre care unii au fugit sau au recidivat în timpul perioadei de cauţiune, iar alţii nu. Pe baza acestor informaţii, vă puteţi imagina că desenaţi un arbore de decizie simplu, cum este cel prezentat în continuare, folosind caracteristicile fiecărui inculpat pentru a construi o diagramă. Odată terminat, arborele de decizie este capabil să prevadă cum s-ar putea comporta un alt inculpat.

Nu trebuie decât să urmăriţi ramurile relevante în funcţie de caracteristicile inculpatului, până când ajungeţi la o predicţie. Atâta vreme cât aceste caracteristici se încadrează în modelul oricui a fugit până în acel moment în timp ce se afla sub cauţiune, predicţia va fi exactă.

---

32 Decision tree (engl.), una dintre metodele de data mining constând în reprezentarea unei serii de reguli care conduce la o valoare sau la o clasă de valori. Poate fi descris şi ca o diagramă care reprezintă variantele de acţiune şi posibilele rezultate pentru fiecare variantă alternativă (n. tr.).



Din acest moment însă arborii de decizie precum aceia pe care i-am construit în şcoală încep să dea greş. Pentru că, evident, nu toată lumea *urmează* tiparul celor care au fugit mai înainte. De sine stătător, arborele acesta va greşi o mulţime de predicţii. Şi asta nu numai din cauză că începem cu un exemplu simplu. Chiar şi cu un set de date uriaş despre cazurile anterioare şi o diagramă îngrozitor de complicată care să îi corespundă, folosirea unui *singur* arbore de decizie ar putea fi doar cu puţin mai bună decât ghicitul la întâmplare.

În cazul în care construieşti mai mult de un arbore, însă, totul se poate schimba. În loc să foloseşti toate informaţiile dintr-odată, există o cale de a dezbina şi de a cuceri. În ceea ce este cunoscut drept un *ensemble*[33], mai întâi, construieşti mii de arbori mai mici din subramificaţii aleatorii de informaţii. Apoi, când ţi se prezintă un nou inculpat, îi ceri pur şi simplu fiecărui arbore să decidă dacă el consideră că solicitarea cauţiunii este sau nu o idee bună. S-ar putea ca nu toţi arborii să fie de acord în unanimitate, iar aceia de sine stătători încă ar putea să emită predicţii nu tocmai valide, dar, dacă iei în considerare media tuturor răspunsurilor, vei putea să-ţi îmbunătăţeşti în mod spectaculos precizia predicţiilor.

Seamănă puţin cu momentul în care adresezi o întrebare publicului, în cadrul emisiunii *Vrei să fii miliardar*? O sală plină de necunoscuţi va avea dreptate mai des decât cea mai inteligentă persoană pe care o cunoşti. (Varianta ajutătoare „întreabă publicul” are o rată de succes de 91%, prin comparaţie cu doar 65% a aceleia

33 „Ansamblu” (fr. în original) (n. tr.).

denumite „Sună un prieten”.) 21 Greşelile făcute de un grup mai numeros se pot anula reciproc, ceea ce are drept rezultat că o mulţime devine mai înţeleaptă decât o singură persoană.

Acelaşi lucru este valabil şi în cazul grupului mare de arbori de decizie, care, luaţi împreună, alcătuiesc o „pădure aleatorie”, *random forest*, (jocul de cuvinte este intenţionat). Pentru că predicţiile algoritmilor se bazează pe tiparele pe care aceştia le învaţă din informaţii, *random forest* este definit ca un algoritm de învăţare automată care apare sub umbrela mai largă a inteligenţei artificiale. (Sintagma „învăţare automată” a apărut pentru prima oară în capitolul „Puterea”, iar, mai departe, vom întâlni mai mulţi algoritmi grupaţi sub această denumire generică; pentru moment, însă, merită menţionat cât de spectaculos sună această descriere atunci când algoritmul este, în esenţă, diagrama pe care obişnuiam s-o desenăm la şcoală la care se adaugă un pic de prelucrare matematică.) *Random forests* s-au dovedit nemaipomenit de utili într-o multitudine de aplicaţii din lumea reală. Sunt utilizaţi de Netflix pentru a prezice ce ţi-ar plăcea să urmăreşti în continuare pe baza preferinţelor tale din trecut22, de Airbnb ca să detecteze conturile frauduloase23 şi în sistemul de sănătate pentru diagnosticarea afecţiunilor (mai multe pe această temă, în capitolul următor).

Atunci când simt folosiţi pentru evaluarea infractorilor, au două imense avantaje faţă de omologii lor umani. În primul rând, orice algoritm va da întotdeauna exact acelaşi răspuns atunci când i se va prezenta acelaşi set de circumstanţe. Consecvenţa este garantată, dar nu cu preţul unei decizii individualizate. Şi mai există un avantaj important: algoritmul va face predicţii mult mai bune.

LUPTA DINTRE OM ŞI MAŞINĂ

În anul 2017, mai mulţi cercetători şi-au propus să descopere cât de adecvate erau predicţiile unui calculator prin comparaţie cu deciziile luate de un anumit număr de

judecători (umani).24

Pentru a-i ajuta în această misiune, cercetătorii aveau acces la dosarele fiecărei persoane arestate în oraşul New York pe o perioadă de cinci ani, între 2008 şi 2013. În acest interval de timp, trei sferturi dintr-un milion de persoane fuseseră audiate pentru a li se stabili cauţiunea, ceea ce a generat cu prisosinţă informaţii pentru a testa performanţele unui algoritm faţă de acelea ale unui judecător uman.

Sistemul judiciar din New York nu utilizase niciun algoritm în aceste cazuri, însă, după ce le-au analizat pe toate, retrospectiv, cercetătorii s-au apucat de treabă şi au construit o mulţime de arbori de decizie pentru a vedea cât de bine ar fi putut cineva să prevadă, când se derulaseră procesele, riscul ca inculpaţii să încalce prevederile eliberării condiţionate. Printre informaţiile despre infractori care au fost introduse se numărau: cazierul, delictul pe care tocmai îl comiseseră şi aşa mai departe. De aici rezulta probabilitatea ca inculpatul să încalce sau nu prevederile eliberării condiţionate.

În evidenţele reale, un număr de 408 283 de inculpaţi fuseseră eliberaţi pe cauţiune chiar înainte de proces. Oricare dintre aceştia era liber să fugă sau să comită alte infracţiuni, ceea ce însemna că puteau fi folosite avantajele unei analize retrospective pentru a se verifica acurateţea predicţiilor algoritmilor şi pe aceea a deciziilor umane, deoarece se ştia cu exactitate cine nu se prezentase ulterior în faţa instanţei (15, 2%) şi cine fusese arestat din nou pentru comiterea altei infracţiuni în perioada eliberării pe cauţiune (25, 8%).

Din păcate pentru ştiinţă, oricărui inculpat considerat de către judecător ca având un risc ridicat i s-ar fi refuzat eliberarea pe cauţiune în acel moment – prin urmare, în acele situaţii nu se putea dovedi în niciun fel dacă evaluările judecătorilor fuseseră corecte sau greşite. Acest lucru complică situaţia. Înseamnă că este imposibil să obţinem o cifră concretă şi exactă care să surprindă cât

de precişi au fost judecătorii per ansamblu. Iar în lipsa unui adevăr întemeiat pe fapte[34], pe observaţii directe ale modului în care s-ar fi comportat inculpaţii care nu fuseseră eliberaţi, nu se poate stabili o acurateţe de ansamblu nici pentru algoritmi. În schimb, trebuie făcute deducţii avizate despre modul în care s-ar fi comportat infractorii rămaşi în închisoare dacă ar fi fost eliberaţi25, iar comparaţiile dintre predicţiile oamenilor şi ale maşinilor trebuie realizate oarecum indirect.

Un lucru este, însă, sigur: judecătorii şi calculatorul nu au emis aceleaşi predicţii. Cercetătorii au arătat că mulţi dintre inculpaţii pe care algoritmul îi semnalase ca fiind băieţi foarte răi fuseseră trataţi de către judecători ca şi când ar fi prezentat un risc scăzut. De fapt, aproape jumătate dintre inculpaţii pe care algoritmul îi semnalase ca reprezentând grupul cu riscul cel mai mare au fost eliberaţi pe cauţiune de judecători.

Dar cine a avut dreptate? Informaţiile au arătat că grupul de care era preocupat algoritmul prezenta, într-adevăr, un risc. Puţin peste 56% din acest grup nu s-au mai prezentat în instanţă, iar 62, 7% au comis noi infracţiuni în perioada eliberării condiţionate – inclusiv pe cele mai abominabile: violul şi crima. Algoritmul prevăzuse toate acestea.

Cercetătorii au susţinut că, oricum ar fi fost folosit, algoritmul lor depăşea cu mult un judecător uman în ceea ce privea performanţele. Iar cifrele le dădeau dreptate. Dacă doreşti să încarcerezi mai puţini oameni în aşteptarea procesului, algoritmul ar putea veni în sprijinul acestui fapt prin condamnarea la închisoare a 41, 8% mai puţini inculpaţi, menţinând totodată aceeaşi rată a criminalităţii. Sau, dacă ai fi mulţumit de procentul actual de inculpaţi eliberaţi pe cauţiune, şi asta ar fi bine: doar fiind mai precis în selectarea inculpaţilor care urmează a fi eliberaţi condiţionat, algoritmul ar putea să reducă procentul celor care fug şi dispar în perioada cauţiunii cu

34 Ground truth (engl.) (vezi glosar, p. 318) (n. red.).

24, 7%.

Aceste beneficii nu sunt doar teoretice. Statul american Rhode Island, unde instanţele au utilizat acest tip de algoritmi în ultimii opt ani, a reuşit să obţină o reducere cu 17% a persoanelor încarcerate şi o scădere de 6% a ratei de recidivă. Aceasta înseamnă că sute de infractori care prezintă un risc scăzut nu mai sunt ţinuţi în închisoare inutil şi că sute de infracţiuni nu au mai fost comise. În plus, dat fiind că detenţia unui prizonier costă peste 30.000 de lire sterline pe an în Marea Britanie26 – pe când în Statele Unite petrecerea unui an într-o închisoare de maximă securitate poate să coste cam cât un an de studiu la Universitatea Harvard27 –, aceasta înseamnă economisirea banilor a sute de mii de contribuabili. Este o situaţie din care toată lumea are de câştigat.

Sau poate că nu?

GĂSIREA LUI DARTH VADER

Fireşte că niciun algoritm nu poate să prevadă în mod *perfect* ce urmează să facă o persoană în viitor. Fiinţele umane sunt prea dezordonate, iraţionale şi impulsive pentru ca vreo prognoză să fie vreodată cu totul *sigură* de ceea ce urmează să se întâmple. Algoritmii ar putea să ofere predicţii mai bune, însă vor continua să facă greşeli. Întrebarea este: ce se întâmplă cu toţi acei oameni al căror punctaj de risc a fost calculat greşit?

Algoritmul poate face două tipuri de greşeli. Richard Berk, profesor de criminologie şi statistică la Universitatea din Pennsylvannia, un pionier în domeniul predicţiei recidivismului, are un mod remarcabil de a le descrie:

Există băieţi buni şi băieţi răi, mi-a spus el. Algoritmul tău întreabă, efectiv: „Care dintre ei sunt Darth Vader-i? Şi care din ei sunt Luke Skywalker-i?”

Să-l laşi liber pe Darth Vader este un tip de eroare cunoscută drept „fals negativ”. Apare ori de câte ori nu reuşeşti să identifici riscul pe care-l prezintă un anumit individ.

Pe de altă parte, să-l închizi pe Luke Skywalker ar fi un „fals pozitiv”. Apare atunci când algoritmul identifică, în mod incorect, un individ ca fiind o persoană cu grad mare de risc.

Aceste două tipuri de erori, fals pozitiv şi fals negativ, nu simt specifice doar fenomenului de recidivă. Le vom întâlni, în mod repetat, pe parcursul acestei cărţi. Orice algoritm care urmăreşte să clasifice ceva se poate face vinovat de comiterea acestor erori.

Algoritmii lui Berk susţin că pot să prezică, şi cu o precizie de 75% încă, dacă cineva urmează să comită o omucidere, ceea ce îi face să fie algoritmii cei mai precişi din câţi există.28 Şi, dacă ne gândim cât de liberă credem noi că este voinţa noastră, acest nivel de precizie este teribil de impresionant. Dar chiar şi o precizie de 75% tot înseamnă că vor exista o mulţime de Luke Skywalker-i cărora le va fi refuzată eliberarea pe cauţiune din închisoare pentru că seamănă cu Darth Vader-ii de afară.

Consecinţele etichetării greşite a unui inculpat devin încă şi mai grave atunci când algoritmii sunt utilizaţi pentru a determina sentinţele, în loc să se limiteze la stabilirea cauţiunii sau a eliberării condiţionate. Aceasta este o realitate modernă: de curând, unele state din SUA au început să le permită judecătorilor să vadă scorul de risc al unui infractor condamnat în timp ce deliberează asupra duratei întemniţării sale. Este o evoluţie care a stârnit o dezbatere aprinsă, şi nu fără motiv: una este să calculezi dacă să eliberezi sau nu o persoană mai devreme din închisoare şi este cu totul altceva să calculezi de la bun început perioada în care aceasta să fie încarcerată.

O parte a problemei este că a decide durata unei sentinţe înseamnă a lua în considerare mult mai mulţi factori decât riscul ca un infractor să recidiveze – ceea ce reprezintă cam tot sprijinul pe care-l poate oferi un algoritm. Un judecător mai trebuie să aprecieze şi riscul pe care infractorul îl reprezintă pentru alţii, efectul de descurajare pe care hotărârea de condamnare îl va avea

asupra altor infractori, problema despăgubirilor acordate victimei şi şansele de reabilitare ale inculpatului. Sunt o mulţime de lucruri de pus în balanţă, aşa încât nu este de mirare că oamenii obiectează împotriva acordării unei importanţe atât de mari unui algoritm în luarea acestui tip de decizie. Şi nici nu este de mirare că oamenilor li se par atât de profund tulburătoare întâmplări precum aceea prin care a trecut Paul Zilly.29

Acesta a fost condamnat pentru furtul unei maşini de tuns iarba. În februarie 2013, s-a prezentat în faţa judecătorului Babler din Baron County, statul Wisconsin, ştiind că avocatul apărării ajunsese deja la o înţelegere cu acuzarea. Ambele părţi căzuseră de acord că, în cazul său, o condamnare pe termen lung nu era cea mai bună soluţie. Zilly se aştepta ca judecătorul să confirme, pur şi simplu, înţelegerea aceasta.

Din păcate pentru el, judecătorii din statul Wisconsin utilizau un algoritm de evaluare a riscului brevetat, denumit COMPAS. Ca şi în cazul instrumentului bugetar din Idaho menţionat la capitolul „Puterea”, funcţionarea internă a COMPAS-ului era considerată un secret comercial. Spre deosebire de instrumentul bugetar, însă, codulalgoritm ului

COMPAS nu este accesibil publicului nici în prezent. Totuşi, din câte ştim, calculele se bazează pe răspunsurile pe care le dă un inculpat atunci când completează un chestionar. Acesta conţine întrebări precum: „Un om flămând are dreptul să fure; eşti de acord sau nu eşti de acord?” şi „Dacă ai locuit împreună cu ambii părinţi, iar aceştia s-au despărţit, ce vârstă aveai în acel moment?” 30 Algoritmul fusese conceput cu unicul scop de a prezice cât de probabil era ca un inculpat să recidiveze în următorii doi ani, iar în această activitate obţinuse un grad de precizie de circa 70%.31 Altfel spus, s-ar fi înşelat, aproximativ, în legătură cu unul din trei inculpaţi. Cu toate acestea, el era folosit de judecători pentru stabilirea sentinţelor.

Punctajul lui Zilly nu era prea bun. Algoritmul îl evaluase ca prezentând risc ridicat de a comite infracţiuni violente în viitor şi risc mediu de recidivă per ansamblu: „Când verific evaluarea riscurilor, spusese judecătorul Babler în instanţă, aceasta vă este cât se poate de nefavorabilă”.

După ce văzuse punctajul lui Zilly, judecătorul a dat dovadă de mai multă încredere în algoritm decât în înţelegerea la care ajunseseră apărarea şi acuzarea, aşa că a respins-o şi i-a dublat sentinţa lui Zilly de la un an petrecut în închisoarea districtuală, la doi într-o închisoare de stat, mai restrictivă.

Este imposibil să ştim cu certitudine dacă Zilly merita punctajul obţinut pentru risc ridicat, cu toate că un procent de acurateţe de 70% pare un prag destul de scăzut pentru a justifica utilizarea algoritmului în vederea respingerii altor factori.

Cazul lui Zilly a fost larg mediatizat, însă nu este singurul exemplu de acest gen. În 2003, Christopher Drew Brooks, un tânăr de 19 ani, a întreţinut relaţii sexuale consensuale cu o fată de 14 ani şi a fost acuzat de întreţinerea unor raporturi sexuale cu un minor de o instanţă de judecată din statul Virginia. Iniţial, instrucţiunile cu privire la sentinţe sugerau o perioadă de încarcerare între şapte şi 16 luni pentru această faptă, însă după ce recomandările au fost modificate pentru a include şi punctajul de risc (care nu fusese calculat de COMPAS, în cazul lui), limita superioară a fost majorată la 24 de luni. Luând în calcul această informaţie, judecătorul l-a condamnat pe Christopher la 18 luni de închisoare.32

Iată unde este problema. Algoritmul folosea vârsta ca factor în calcularea punctajului de recidivă. Faptul că fusese condamnat pentru un delict sexual la o vârstă atât de fragedă era în defavoarea lui Brooks, chiar dacă asta însemna totuşi că vârsta îi era mai apropiată de cea a victimei. De fapt, dacă Brooks ar fi avut 36 de ani (prin urmare, cu 22 de ani mai mult decât fata), algoritmul ar fi

recomandat să nu fie trimis deloc la închisoare! 33

Acestea nu sunt singurele exemple în care oamenii au avut mai multă încredere în rezultatele obţinute de un calculator decât în propria putere de judecată şi nu vor fi nici ultimele. Se pune întrebarea: ce este de făcut în această privinţă? Curtea Supremă din Wisconsin vine cu o sugestie. Referindu-se, în mod special, la pericolul ca judecătorii să se bazeze prea mult pe algoritmul COMPAS, aceasta a decretat:

Ne aşteptăm ca tribunalele să manifeste discernământ, în privinţa fiecărui inculpat în parte, atunci când vor evalua vreunul dintre punctajele de risc obţinute cu ajutorul lui COMPAS.34

Richard Berk sugerează însă că aceasta ar putea fi o abordare prea optimistă:

Judecătoriile sunt preocupate să nu facă greşeli – în special, judecătorii numiţi din oficiu. Algoritmul le oferă o modalitate de a-şi micşora volumul de muncă şi, totodată, de a nu fi consideraţi răspunzători pentru o anumită sentinţă.35

Dar mai este o problemă. Dacă un algoritm clasifică o persoană ca prezentând risc ridicat, iar judecătorul refuză, în consecinţă, punerea ei în libertate, nu există nicio cale de a şti cu certitudine dacă algoritmul i-a analizat perspectivele cu acurateţe. Să-l luăm drept exemplu pe Zilly. Poate că ar fi continuat să aibă un comportament violent. Poate că nu. Poate că a fi considerat un condamnat cu nivel de risc ridicat şi a fi trimis într-o închisoare de stat nu au făcut decât să-i impună o cu totul altă traiectorie în viaţă decât dacă s-ar fi ţinut cont de înţelegerea la care se ajunsese la tribunal. Neavând nicio modalitate de a verifica predicţiile algoritmilor, nu putem şti dacă judecătorul a avut dreptate să dea crezare punctajului pentru risc şi nu avem nicio cale de a verifica dacă Zilly era, de fapt, un Darth Vader sau un Luke Skywalker.

Această problemă nu are o soluţie facilă. Cum îi convingi pe oameni să apeleze la o doză sănătoasă de bun-

simţ când se pune problema utilizării acestor algoritmi? Şi, chiar dacă acest lucru ar fi posibil, mai apare o problemă legată de predicţia recidivei. Fără îndoială, cea mai controversată dintre toate.

PREJUDECĂŢILE MAŞINILOR

În anul 2016, redacţia publicaţiei independente on-line ProPublica, prima care a relatat cazul lui Zilly, a studiat în detaliu algoritmul COMPAS şi a reconstituit în sens invers predicţiile pe care le făcuse acesta despre perspectivele de recidivare a peste 7.000 de infractori reali din Florida36, în cazuri datând din 2013 până în 2014. Cercetătorii doreau să verifice acurateţea punctajelor obţinute de COMPAS, urmărind care dintre infractori recidivaseră efectiv. Mai doreau, de asemenea, să vadă dacă exista vreo diferenţă între riscurile prezise pentru infractorii de culoare şi cei albi.

Cu toate că algoritmul nu include, în mod explicit, rasa ca factor, jurnaliştii au descoperit că nu toţi inculpaţii erau trataţi în mod egal la efectuarea calculelor. Deşi şansele de eroare ale algoritmului erau, în linii mari, aceleaşi pentru infractorii de culoare şi pentru cei albi, el făcea diverse tipuri de erori în cazul fiecărui grup rasial în parte.

Dacă erai unul dintre infractorii care nu mai comiseseră nicio faptă reprobabilă după arestarea iniţială, prin urmare, un adevărat Luke Skywalker, algoritmul era de două ori mai predispus să te eticheteze în mod greşit ca având un risc ridicat dacă erai o persoană de culoare, decât dacă erai alb. Rezultatele fals pozitive ale algoritmului erau în mod disproporţionat atribuite persoanelor de culoare. În schimb, printre infractorii care au comis o nouă infracţiune în următorii doi ani, printre Darth Vadder-i, adică, era de două ori mai probabil ca deţinuţii albi să fie evaluaţi de algoritm ca prezentând un risc scăzut pentru societate, faţă de deţinuţii de culoare. Rezultatele fals negative ale algoritmului erau, în mod disproporţionat, atribuite albilor.

Deloc surprinzător, analiza făcută de ProPublica a stârnit indignare pe tot cuprinsul Statelor Unite şi chiar mai departe. S-au scris sute de articole care exprimau dezaprobarea fermă a autorilor faţă de utilizarea unor calcule impersonale în procesul uman de a face dreptate şi se condamna folosirea unor algoritmi imperfecţi, părtinitori, în decizii care aveau un impact atât de mare asupra viitorului unei persoane. Multe dintre aceste critici sunt greu de contrazis – toată lumea merită să aibă parte de un tratament corect şi egal indiferent cine evaluează cazul respectiv –, iar studiul realizat de ProPublica nu pune algoritmul într-o lumină prea bună.

Să fim, însă, atenţi, pentru o clipă, la tendinţa noastră de a înlătura „algoritmii imperfecţi”. Înainte de a respinge cu totul utilizarea acestora în sistemul judiciar, merită să ne punem întrebarea: cum v-aţi aştepta să fie un algoritm nepărtinitor?

Aţi dori, desigur, ca acesta să facă predicţii la fel de corecte pentru condamnaţii de culoare şi pentru cei albi. Pare, de asemenea, rezonabil să ceri ca toţi factorii ce se consideră a reprezenta un „risc ridicat” să fie la fel pentru toată lumea. Algoritmul ar trebui să fie la fel de performant în selectarea infractorilor predispuşi să recidiveze indiferent de grupul rasial (sau oricare altul) căruia îi aparţin aceştia. La care, aşa cum a subliniat ProPublica, se adaugă faptul că algoritmul ar trebui să facă acelaşi tip de greşeli cu aceeaşi frecvenţă pentru toată lumea – indiferent de rasă.

Niciuna dintre aceste patru afirmaţii nu pare să fie din calea-afară de ambiţioasă. Numai că apare, totuşi, o problemă. Din nefericire, unele tipuri de corectitudine sunt, din punct de vedere matematic, incompatibile cu altele.

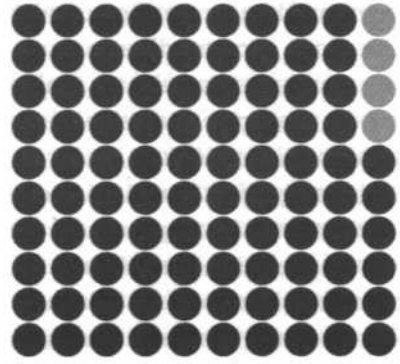
Daţi-mi voie să vă lămuresc. Imaginaţi-vă că opriţi oamenii pe stradă şi utilizaţi un algoritm care prezice dacă o anumită persoană urmează să comită un omor. Însă, din moment ce marea majoritate a omorurilor sunt comise de

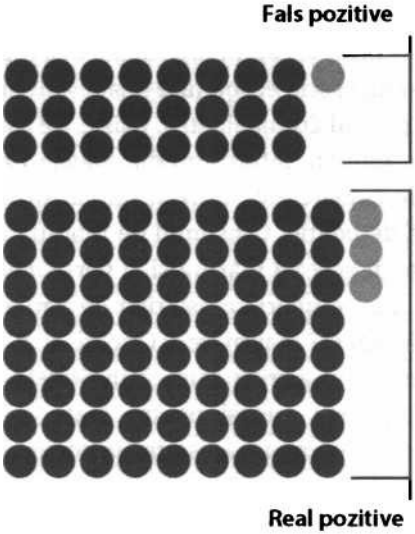
bărbaţi (în întreaga lume, 96% dintre ucigaşi sunt bărbaţi) 37, dacă algoritmul care depistează ucigaşi trebuie să facă predicţii exacte, el va identifica, în mod necesar, mai mulţi bărbaţi decât femei ca prezentând un risc mai mare de a comite acest gen de faptă.

Să presupunem că algoritmul nostru de detectare a ucigaşilor are o rată de predicţie de 75%. Asta înseamnă că trei sferturi din persoanele pe care algoritmul le consideră ca având un risc crescut sunt, într-adevăr, nişte Darth Vader-i.

În cele din urmă, după ce aţi oprit destui necunoscuţi, veţi avea 100 de persoane semnalate de algoritm ca fiind potenţiali ucigaşi. Pentru a face corelaţia cu statistica făptaşilor, 96 dintre aceşti 100 vor fi, obligatoriu, bărbaţi. Patru vor fi femei, în continuare este o poză pentru ilustrare. Bărbaţii sunt reprezentaţi de cerculeţele închise la culoare, femeile apar ca nişte cerculeţe de un gri-deschis.

Acum, din moment ce algoritmul face predicţii corecte atât pentru bărbaţi, cât şi pentru femei într-o proporţie egală de 75%, un sfert dintre femei şi un sfert din bărbaţi vor fi, de fapt, Luke Skywalker-i: persoane care au fost incorect identificate ca având un nivel de risc ridicat, atunci când, de fapt, ele nu prezintă niciunul.





Când faci calculele, aşa cum se poate vedea din a doua imagine de mai sus, vor fi acuzaţi pe nedrept mai mulţi bărbaţi nevinovaţi decât femei nevinovate doar în virtutea faptului că bărbaţii săvârşesc mai multe omoruri decât femeile.

Asta nu are nimic de-a face nici cu crima în sine şi nici cu algoritmul: este doar o certitudine matematică. Rezultatul este viciat de prejudecată pentru că realitatea însăşi este deformată într-un mod care întăreşte această prejudecată. Mai mulţi bărbaţi comit omucideri, aşadar, mai mulţi bărbaţi vor fi acuzaţi, în mod greşit, că prezintă

riscul de a ucide.[35]

Dacă nu se întâmplă ca procentul persoanelor care comit crime să fie acelaşi în fiecare grup de infractori, este matematic imposibil să concepi un test care, în acelaşi timp, să facă predicţii tot atât de exacte pe toată linia şi să eticheteze persoanele din fiecare grup de infractori drept fals pozitive şi fals negative în aceeaşi proporţie.

Desigur, persoanele de culoare din SUA au făcut obiectul unor prejudecăţi şi nedreptăţi cumplite de-a lungul secolelor. Din această cauză, ele continuă să apară disproporţionat de des la baza piramidei socio-economice şi la nivelurile superioare ale statisticilor privind criminalitatea. Există, de asemenea, dovezi că – pentru anumite infracţiuni din Statele Unite cel puţin – populaţia de culoare este într-un număr disproporţionat de mare vizată de poliţie. Consumul de marijuana, de exemplu, se întâlneşte în aceeaşi măsură şi la persoanele de culoare şi la cele albe, însă numărul arestărilor de afro-americani poate fi chiar şi de opt ori mai mare.38 Oricare ar fi motivul acestei discrepanţe, tristul adevăr este că în Statele Unite numărul arestărilor nu este acelaşi pentru toate grupurile rasiale. Persoanele de culoare sunt arestate din nou mai des decât albii. Algoritmul nu le discerne după culoarea pielii, ci după efectele mult prea previzibile ale vieţuirii în societatea americană, care, din motive istorice, este atât de dezechilibrată. Până când toate grupurile rasiale vor ajunge să fie arestate în proporţii egale, acest gen de inechitate va rămâne, în continuare, o certitudine matematică.

Asta nu înseamnă însă nici că ar trebui să respingem cu totul articolul din ProPublica. Analiza pe care o face acesta subliniază cât de uşor pot algoritmii să perpetueze inechităţile trecutului. Nu înseamnă însă nici că ar fi cazul

35 Un rezultat ca acesta poate surveni chiar dacă nu este folosit, în mod explicit, genul cuiva drept factor al algoritmului. Atât timp cât predicţia se bazează pe factori care se corelează mai mult cu un grup decât cu altul (cum ar fi antecedentele de manifestări violente ale acuzatului), acest tip de inechitate poate să apară (n. a.).

să găsim scuze algoritmului COMPAS. Orice companie care obţine un profit din analizarea informaţiilor despre diverse persoane are responsabilitatea morală (dacă nu cumva şi legală) să-şi recunoască defectele şi punctele vulnerabile. Equivant (fosta Northpointe), în schimb, firma care produce COMPAS, continuă să considere configuraţia algoritmului drept un secret care trebuie păzit îndeaproape pentru a-şi proteja drepturile de proprietate intelectuală asupra acestuia.39

Avem totuşi câteva opţiuni. Nu există nimic intrinsec în aceşti algoritmi care să îi determine să repete prejudecăţile trecutului. Totul se rezumă la informaţiile care le sunt furnizate. Putem alege să fim „adepţii unui empirism cras" (cum spune Richard Berk) şi să ne luăm, în continuare, după cifrele deja existente sau putem decide că statu-quo-ul este nedrept şi, în consecinţă, să modificăm numerele statisticilor.

Ca să ofer o analogie, o să vă rog să încercaţi să daţi o căutare pe Google Imagini pentru „profesor de matematică". Poate că nu veţi fi surprinşi să descoperiţi că majoritatea imaginilor reprezintă bărbaţi albi, de vârstă mijlocie, stând în picioare în faţa unor table negre, pline de însemne făcute cu creta. Căutarea mea a întors numai o singură imagine cu o femeie în primele 20 de imagini, ceea ce este o reflectare deprimant de corectă a realităţii: în jur de 94% dintre profesorii de matematică sunt bărbaţi.40 Dar, oricât de precise ar fi rezultatele, se poate afirma că să utilizezi algoritmii doar ca pe o oglindă menită să reflecte lumea reală nu este întotdeauna util, mai ales atunci când oglinda reflectă o realitate prezentă care există numai din cauza unor îndelungate secole de nedreptăţi. Dacă ar dori.

Google ar putea să-şi ajusteze subtil algoritmul, astfel încât să redea, în mod prioritar, imagini cu femei sau cu profesori care să fie de altă rasă, nu doar din cea albă, pentru a echilibra astfel puţin balanţa şi a ajunge să reflecte societatea spre care tindem, mai degrabă decât pe

aceea în care trăim.

Acelaşi lucru se întâmplă şi în sistemul judiciar. Atunci când utilizăm un algoritm, s-ar cuveni să ne întrebăm, efectiv: ce procent dintr-un anumit grup ne-am aştepta să prezinte un risc ridicat într-o societate perfect echitabilă? Algoritmul ne oferă opţiunea să sărim direct la cifra aceea. Sau, dacă decidem că îndepărtarea dintr-odată a tuturor prejudecăţilor din sistemul judiciar nu este oportună, am putea cere, în schimb, algoritmului să acţioneze progresiv, în timp, spre îndeplinirea acestui deziderat.

Dispunem de opţiuni şi în ceea ce priveşte modul în care îi tratăm pe inculpaţii care prezintă un risc ridicat la eliberarea pe cauţiune, unde riscul ca un inculpat să nu-şi mai facă apariţia la următorul termen al procesului este o componentă principală a predicţiei algoritmului, abordarea standard este aceea de a refuza acordarea cauţiunii oricărei persoane care prezintă un grad ridicat de risc. Dar algoritmul ar putea oferi şi prilejul de a afla de ce o persoană ar putea să nu se prezinte la o înfăţişare la tribunal. Are acces la un mijloc de transport adecvat ca să ajungă acolo? Are probleme cu îngrijirea copiilor pe care nu îi poate lăsa nimănui în grijă, ceea ce o s-o împiedice să se prezinte în instanţă? Există nedreptăţi sociale pe care algoritmul ar putea fi programat să le atenueze, în loc să le exacerbeze?

Răspunsurile la aceste întrebări ar trebui să vină din partea forumurilor deschise de dezbateri publice şi din sferele guvernării, mai degrabă decât din sălile de şedinţă ale companiilor private. Din fericire, pentru a controla această industrie se solicită din ce în ce mai insistent apariţia unui organism de reglementare a algoritmilor. Aşa cum procedează US Food and

Drug Administration[36] în cazul produselor

36 Cunoscută şi cu acronimul FDA, este una dintre agenţiile federale din Statele Unite din domeniul medical. Este responsabilă cu reglementarea produselor alimentare (destinate oamenilor şi animalelor), a suplimentelor dietetice, medicamentelor, cosmeticelor,

farmaceutice, acest organism ar putea testa, în spatele uşilor închise, acurateţea, consecvenţa şi distorsiunile algoritmilor şi ar avea autoritatea de a aproba sau de a respinge utilizarea unui produs care ar urma să afecteze fiinţele umane reale. Până atunci, însă, este extraordinar de important ca organizaţii cum este ProPublica să continue să tragă la răspundere companiile care concep algoritmi, atât timp cât acuzaţiile că aceştia au fost părtinitori nu se transformă în solicitări de a-i interzice complet, cel puţin, nu fără a cerceta cu atenţie ce ne-ar mai rămâne în acest caz.

DECIZII DIFICILE

Există un aspect foarte important la care trebuie să mă refer. Dacă respingem algoritmii, cu ce fel de sistem judiciar am mai rămâne? Pentru că inconsecvenţa nu este singurul defect de care s-a dovedit că se fac vinovaţi judecătorii.

Din punct de vedere legal, rasa, genul şi clasa socială a unui inculpat nu ar trebui să influenţeze niciodată nicio hotărâre judecătorească. (în definitiv, se presupune că justiţia este oarbă, nu?) Cu toate acestea, cei mai mulţi dintre judecători doresc să fie cât mai nepărtinitori cu putinţă, dovezile au arătat în mod repetat că judecătorii, efectiv, discriminează. Studii din Statele Unite au dovedit că acuzaţii de culoare vor fi trimişi la închisoare, în medie, un timp mai îndelungat41, că au şanse mai mici să li se acorde eliberarea pe cauţiune42, dar şanse mai mari să fie condamnaţi la moarte43 şi, odată ce se află pe lista condamnaţilor la moarte, există o mai probabilitate să fie executaţi.44 Alte cercetări au arătat că, pentru aceeaşi infracţiune, bărbaţii sunt trataţi cu mai multă severitate decât femeile45 şi că inculpaţilor cu niveluri de venit şi de educaţie scăzute li se dau condamnări mai lungi.46

Ca şi în cazul algoritmilor, nu atât prejudecăţile propriu-zise generează aceste rezultate părtinitoare, cât faptul că istoria se repetă. Prejudecăţile sociale şi

---

dispozitivelor medicale şi a acelora care emit radiaţii etc. (n. tr.).

culturale pot să apară, pur şi simplu, ca o consecinţă automată a modului în care oamenii iau deciziile.

Pentru a explica de ce se întâmplă acest lucru, trebuie să înţelegem, mai întâi, câteva lucruri simple despre intuiţia umană, aşa că vom părăsi, pentru moment, sala de judecată, cât timp reflectaţi la următoarea problemă:

O bâtă de baseball şi o minge costă împreună 1, 10 lire.

Bâta costă cu o liră mai mult decât mingea.

Cât costă mingea?

Această problemă, care a fost introdusă în best-sellerul economistului şi psihologului Daniel Kahneman, laureat al Premiului Nobel pentru Economie, *Gândire rapidă, gândire lentă*[37] *7*, ilustrează o capcană importantă în care cădem cu toţii atunci când gândim.

Întrebarea are un răspuns incorect, uşor de sesizat la o gândire mai atentă, dar şi un răspuns incorect, care se iveşte imediat în minte. Răspunsul la care v-aţi gândit, mai întâi, a fost, bineînţeles, 10 pence.*

Să nu vă pară rău dacă nu aţi găsit răspunsul corect (5 penny). Nu l-au găsit nici 71, 8% din judecători atunci când au fost întrebaţi.48 Chiar şi cei care, până la urmă, reuşesc să găsească răspunsul corect, trebuie să-şi reprime, pe moment, impulsul de a-şi urma intuiţia.

Competiţia dintre intuiţie şi gândirea chibzuită este esenţială în cazul relatării noastre despre hotărârile judecătoreşti. Psihologii sunt de acord, în linii mari, că dispunem de două moduri de gândire. Primul sistem este automat, instinctiv, dar predispus la greşeli. (Este sistemul responsabil de răspunsul „10 penny" care ni s-a ivit în minte la întrebările de mai sus.) Al doilea sistem este încet, analitic, chibzuit, însă, adesea, destul de leneş.49

Dacă întrebi o persoană cum a ajuns să ia o decizie, al doilea sistem va formula un răspuns, dar, aşa cum spune

37 Dacă mingea ar costa 10 penny ar însemna că bâta a costat 1,10 lire, iar preţurile celor două produse, însumate, ar fi 1,20 lire (n. a.).

Daniel Kahneman: „acesta vine adesea în sprijinul sau justifică ideile şi sentimentele generate de primul sistem”.50

În această situaţie, judecătorii nu diferă de noi, ceilalţi. În definitiv, şi ei sunt oameni, cu aceleaşi toane şi slăbiciuni pe care le avem cu toţii. Adevărul este că minţile noastre pur şi simplu nu sunt echipate pentru evaluarea solidă şi raţională a problemelor mari şi complicate. Nu putem cântări cu uşurinţă diverşii factori ai unei situaţii şi să legăm totul de o manieră logică, în timp ce împiedicăm primul sistem intuitiv să intervină şi să o ia pe tot soiul de scurtături cognitive.

Când ne referim la justiţie, de exemplu, s-ar putea crede că judecătorii sunt capabili să evalueze un caz în totalitate, cântărind cu atenţie argumentele pro şi contra înainte de a lua o hotărâre. Din păcate, dovezile indică altceva. Psihologii au arătat, în schimb, că toată strategia folosită de judecători se rezumă la a trece în revistă o listă de semnale de avertizare pe care şi-au fixat-o în memorie. Dacă relatarea acuzatului îi evocă oricare dintre aceste semnale – condamnări din trecut, absenţa legăturilor cu comunitatea, solicitarea procurorului –, judecătorul se va opri din chibzuire şi îi va refuza eliberarea pe cauţiune.51

Problema este că foarte multe dintre aceste semnale de avertizare au legătură cu rasa, genul şi nivelul de educaţie al inculpatului. Judecătorii nu se pot abţine să nu se bazeze pe intuiţie mai mult decât ar trebui; iar procedând în felul acesta, ei perpetuează fără să vrea inechităţile sistemului.

Şi asta nu-i tot! Din păcate, abia dacă am atins o mică parte a problemei, abia dacă am deschis discuţia despre cât de jalnici, cât de incapabili sunt oamenii să se comporte ca nişte judecători corecţi şi nepărtinitori.

Dacă v-aţi convins vreodată singuri că un anumit obiect de îmbrăcăminte extrem de costisitor a devenit o afacere grozavă doar pentru că era vândut cu 50% mai ieftin (aşa cum eu, una, fac de obicei), atunci ştiţi tot ce e

de ştiut despre aşa-numitul „efect de ancorare”. Acesta denumeşte faptul că ne este greu să stabilim o valoare numerică pentru diverse lucruri şi găsim că este mult mai simplu să facem comparaţii între diverse valori, decât să găsim din senin o singură valoare. Comercianţii folosesc efectul de ancorare de ani de zile nu numai ca să influenţeze cât de mult apreciem anumite produse, dar şi ca să controleze cantitatea de produse pe care o cumpărăm. Gândiţi-vă, spre exemplu, la acele etichete indicatoare din supermarketuri pe care scrie: „Nu pot fi cumpărate mai mult de 12 cutii cu supă de către un client”. Ele nu sunt menite să împiedice iubitorii de supă să epuizeze stocul de produse, aşa cum aţi putea crede. Dimpotrivă, se află acolo pentru a vă manipula, în mod subtil, percepţia asupra numărului de cutii cu supă de care aveţi nevoie. Creierul se ancorează de numărul 12 şi ajustează în jos. Iar un studiu din anii 1990 a arătat că un astfel de semnal ar putea să crească media vânzărilor per client de la 3, 3 cutii cu supă, la 7.52

Acum nu veţi mai fi deloc surprinşi să aflaţi că şi judecătorii sunt sensibili la efectul de ancorare. Ei sunt mai predispuşi să ceară despăgubiri mai însemnate dacă procurorul solicită o sumă mai mare53 şi să pronunţe o condamnare mai lungă dacă procurorul cere o pedeapsă mai severă.54 Un studiu a arătat chiar că s-ar putea influenţa în mod considerabil durata unei condamnări, într-un caz ipotetic, dacă ar fi trimis un jurnalist la judecător în timpul unei pauze, iar acesta ar introduce, cu subtililitate, în conversaţie sugerarea unei anumite sentinţe.

(» Credeţi că sentinţa, în cazul de faţă, ar trebui să fie condamnarea pe o durată mai mare sau mică de trei ani?”) 55 Probabil că lucrul cel mai îngrozitor este că poţi influenţa decizia unui judecător dacă îi ceri, pur şi simplu, să dea cu banul înainte de a intra în sala de judecată.56 Până şi judecătorii cu cea mai mare experienţă au fost susceptibili la acest gen de manipulare.57

Şi mai există o fisură în modul în care fiinţele umane

fac comparaţii între numere, între cantităţi, care are un anumit impact asupra caracterului echitabil al pronunţării sentinţei. Poate aţi remarcat deja această ciudăţenie a minţii omeneşti. Atunci când muzica e dată tare, creşterea volumului cu o unitate este mai puţin percepută decât aceeaşi creştere absolută la un volum mai mic; o creştere a preţului de la 1, 20 lire sterline la 2, 20 pare exagerat, însă o creştere de la 67 de lire sterline la 68 nu prea contează. Iar timpul pare să treacă mai repede pe măsură ce înaintăm în vârstă. Toate acestea se întâmplă pentru că la fiinţele umane simţurile funcţionează mai mult în termeni relativi decât în valori absolute. Nu percepem fiecare an ca pe un interval de timp fix, ci îl sesizăm ca pe o fracţiune din ce în ce mai scurtă din viaţa pe care am trăit-o. Dimensiunea fragmentelor de timp, bani sau volum sonor pe care le percepem urmează o expresie matematică foarte simplă, cunoscută drept Legea lui Weber.

Pe scurt, Legea lui Weber afirmă că întotdeauna cea mai mică modificare a stimulului care poate fi percepută, aşa-numita „cea mai mică diferenţă perceptibilă”, este proporţională cu stimulul iniţial. Deloc surprinzător, această descoperire a fost exploatată şi de comercianţi. Ei ştiu exact cât de mult pot să micşoreze o tabletă de ciocolată fără ca niciunul dintre clienţi să bage de seamă sau ştiu precis cât de mult pot să majoreze preţul unui produs înainte ca vreunul dintre noi să considere că ar fi mai bine să-l caute şi în alt magazin.

În contextul justiţiei, problema este că Legea lui Weber influenţează durata condamnării alese de judecători. Diferenţele dintre sentinţe sunt tot mai mari, pe măsură ce pedepsele devin mai severe. Dacă o infracţiune este cu puţin mai gravă decât una care merită o condamnare de 20 de ani, atunci trei luni în plus, să spunem, nu par suficiente: parcă nu ar exista o diferenţă suficient de mare între o durată a sentinţei de 20 de ani şi una de 20 de ani şi trei luni. Dar bineînţeles că există: trei luni de închisoare sunt, totuşi, trei luni de închisoare,

indiferent de ce a fost înaintea lor. Cu toate acestea, în loc să mai adauge doar câteva luni, judecătorii vor trece la următoarea diferenţă perceptibilă a duratei sentinţei, care, în cazul de faţă, este de 25 de ani.58 Ştim că acest lucru se întâmplă pentru că putem să comparăm duratele sentinţelor pronunţate în realitate cu cele pe care le-ar prezice Legea lui Weber. Un studiu din 2017 a analizat peste 100.000 de sentinţe, atât din Marea Britanie, cât şi din Australia, şi a găsit că până la 99% dintre inculpaţii consideraţi vinovaţi au primit o condamnare care corespundea acestui enunţ.59

Mandeep Dhami, autorul principal al studiului, mi-a declarat:

Nu are importanţă ce tip de infracţiune ai comis sau ce fel de inculpat eşti, nici în ce ţară ai fost condamnat şi nici chiar dacă ai fost condamnat la închisoare sau la prestarea unor ore de muncă în beneficiul societăţii.

Tot ce conta era numărul care se ivea în mintea judecătorului şi care i se părea acestuia cât de cât corect.

Iar dacă ar fi să mă refer la cât de părtinitori pot fi judecătorii, din păcate, mi-ar fi posibil să continui la nesfârşit! Judecătorii care au fiice sunt mai predispuşi să ia decizii favorabile femeilor.60 Judecătorii simt mai puţin dispuşi să ofere eliberare pe cauţiune atunci când echipa sportivă locală a pierdut recent un meci. Iar un studiu celebru sugerează chiar că până şi momentul din zi îţi afectează şansele de a obţine un rezultat favorabil în instanţă.61 Deşi studiul acesta încă nu a fost reprodus62 şi sunt în curs câteva dezbateri în privinţa amplorii efectului, se pare că ar exista dovezi că susţinerea unei depoziţii chiar înainte de masa de prânz poate fi un dezavantaj: judecătorii din studiul iniţial erau mai dispuşi să ofere eliberarea pe cauţiune dacă tocmai se întorseseră după o pauză şi cel mai puţin dispuşi atunci când se apropia pauza de masă.

Un alt studiu a arătat că un singur judecător va evita să pronunţe prea multe sentinţe asemănătoare la rând.

Prin urmare, şansele de a ţi se oferi eliberarea pe cauţiune scad ameţitor dacă patru cazuri care au fost judecate chiar înaintea ta au avut succes în această privinţă.63

Unii cercetători mai susţin că până şi modul în care percepem persoanele necunoscute se schimbă în funcţie de temperatura băuturii pe care o ţinem în mână. Astfel, dacă ne-a fost oferită o băutură caldă chiar înainte de a întâlni o persoană nouă, sugerează ei, suntem mai dispuşi să considerăm că aceasta are o personalitate mai caldă, mai generoasă, mai plină de afecţiune.6*

lista aceasta lungă conţine numai chestiunile care pot fi măsurate. Fără îndoială că există nenumăraţi alţi factori care ne influenţează subtil comportamentul şi care nu pot fi testaţi într-o sală de judecată.

ÎN CONCLUZIE

Voi fi sinceră cu dumneavoastră. Când am aflat pentru prima oară că algoritmii erau utilizaţi în instanţele de judecată, m-am opus acestei idei. Un algoritm va face greşeli, iar când o eroare poate însemna că un om îşi va pierde dreptul la libertate, nu mi se părea un gest responsabil să laşi această putere la latitudinea unei maşini.

Nu sunt singura care gândeşte astfel. Multe persoane (poate chiar majoritatea) care ajung să fie la cheremul sistemului de justiţie penală sunt de aceeaşi părere. Mandeep Dhami mi-a spus ce simţeau inculpaţii cu care lucrase faţă de modul în care erau luate deciziile privitoare la viitorul lor:

Chiar dacă ştiau că judecătorul ar fi putut face mai multe greşeli, inculpaţii continuau să prefere să fie judecaţi de o fiinţă umană, nu de un algoritm. Îşi doreau să existe o relaţie umană.

La fel şi avocaţii, dacă tot a venit vorba. Un avocat al apărării care îşi practică meseria la Londra, cu care am vorbit, mi-a spus că rolul său în sala de judecată era să exploateze incertitudinea din sistem, ceva ce algoritmii vor face mai dificil să se întâmple. „Cu cât sentinţele devin mai

previzibile, cu atât rămâne mai puţin loc pentru arta pledoariei”.

Când, însă, am întrebat-o pe Mandeep Dhami cum s-ar simţi dacă ar fi ea însăşi ameninţată cu închisoarea, a dat un răspuns complet opus:

Nu vreau ca o persoană să-şi folosească intuiţia atunci când ia o hotărâre în privinţa viitorului meu. Vreau să apeleze la o strategie bine gândită. Dorim să păstrăm independenţa judiciară şi bunul-plac al judecătorilor de parcă ar fi ceva sacrosanct! De parcă ar fi ceva nespus de bun! Chiar dacă toate cercetările arată că nu este aşa. Independenţa judiciară nu este deloc un lucru nemaipomenit.

Ca mai toată lumea, şi eu consider că deciziile judecătorilor ar trebui să fie cât mai nepărtinitoare cu putinţă. Aceştia ar trebui să se orienteze după faptele săvârşite de o persoană, nu după grupul căruia se poate întâmpla să îi aparţină. Din acest punct de vedere, algoritmii nu prea sunt suficient de buni. Însă nu este de ajuns doar să indici ce este greşit la aceştia. Alegerea nu se face între un algoritm cu imperfecţiuni şi un sistem imaginar perfect. Singura comparaţie cinstită care poate fi făcută este între un algoritm şi ceea ce ne-ar rămâne în absenţa sa.

Cu cât m-am informat mai mult, cu cât am vorbit cu mai multe persoane, cu atât am ajuns să cred că ne aşteptăm la prea mult de la judecătorii noştri, care nu sunt decât nişte biete fiinţe din carne şi oase. Inechitatea este parte integrată a tuturor sistemelor noastre umane. Pentru fiecare Christopher Drew Brooks tratat incorect de un algoritm, există nenumărate cazuri, ca acela al lui Nicholas Robinson, în care un judecător greşeşte pe cont propriu. Dar, dacă există un algoritm – chiar şi unul imperfect – care lucrează cu judecătorii pentru a veni în sprijinul gândirii lor adesea defectuoase, acesta reprezintă, cred eu, un pas în direcţia corectă. Pentru că un algoritm bine conceput şi calibrat aşa cum trebuie poate ajuta la

eliminarea tendenţiozităţii sistematice şi a erorilor aleatorii. Nu poţi schimba o întreagă cohortă de judecători, mai ales dacă nici ei înşişi nu sunt în stare să-ţi spună cum au luat deciziile pe care le-au luat.

Pentru a crea un algoritm care să fie utilizat în sistemul judiciar trebuie să ne oprim însă şi să ne gândim temeinic care este, mai exact, scopul acestui sistem. În loc să închidem ochii şi să sperăm că totul va fi bine, algoritmii ne cer să ne formăm o idee clară şi lipsită de ambiguitate despre ce vrem ca ei să realizeze negreşit, precum şi o înţelegere fermă a eşecurilor umane pe care vin să le înlocuiască. Silindu-ne, astfel, să purtăm o dezbatere dificilă despre cum ar trebui luate deciziile într-un tribunal. Nu va fi simplu, dar aceasta este cheia pentru a stabili dacă algoritmii vor putea fi vreodată suficient de buni în acest scop.

În sistemul judiciar se manifestă tensiuni care tulbură apele şi fac ca acest gen de întrebări să-şi găsească doar cu mare dificultate răspunsurile. Dar există şi alte domenii în care algoritmii pătrund treptat şi în care luarea deciziilor este mult mai puţin conflictuală, iar contribuţia obiectivă şi pozitivă pe care o pot aduce algoritmii în societate poate fi mult mai clar definită.

Capitolul 4 MEDICINA

În anul 2015, nişte oameni de ştiinţă animaţi de spiritul pionieratului au realizat un studiu neobişnuit pentru a determina acurateţea modului în care era diagnosticat cancerul mamar.1 Ei le-au pus la dispoziţie câte un monitor cu ecran tactil (*touch-screen*) celor 16 participanţi la test şi le-au cerut să trieze imaginile cu ţesut mamar. Probele fuseseră prelevate de la femei reale al căror ţesut mamar fusese îndepărtat prin biopsie, secţionat în felii subţiri şi colorat cu diverse substanţe chimice pentru a evidenţia vasele de sânge şi canalele mamare în nuanţe de roşu, purpuriu şi albastru. Participanţii la test nu trebuiau decât să decidă dacă tiparele din imagini sugerau că printre celulele sănătoase

se ascundeau şi unele canceroase.

După o scurtă perioadă de antrenament, participanţii la test au fost puşi la lucru şi au dat rezultate impresionante. Lucrând independent, au evaluat corect 85% dintre probe.

Atunci cercetătorii au descoperit un lucru remarcabil. Dacă începeau să cumuleze răspunsurile – combinând voturile fiecărui participant la test pentru a obţine o evaluare generală a imaginii –, rata de corectitudine creştea brusc, ajungând la 99%!

Cu adevărat extraordinară la acest studiu nu era doar priceperea participanţilor la test. Era identitatea lor. Aceşti salvatori curajoşi nu erau oncologi. Nu erau nici patologi. Nu erau nici asistente medicale. Nu erau nici măcar studenţi la Medicină. Erau doar nişte porumbei.

Locurile de muncă ale patologilor sunt încă în siguranţă, cel puţin, pentru o vreme – cred că nici măcar oamenii de ştiinţă care au pus la punct studiul nu ar îndrăzni să sugereze că medicii ar trebui înlocuiţi cu banalii şi bine-cunoscuţii porumbei. Însă experimentul a demonstrat un aspect important: descoperirea tiparelor ascunse printre grupurile de celule nu este o abilitate specific umană. Aşadar, dacă un porumbel se poate descurca onorabil în rezolvarea unei asemenea sarcini, de ce n-ar putea s-o facă şi un algoritm?

VÂNĂTORII DE TIPARE

Întreaga istorie a medicinei moderne şi toată practicarea sa se bazează pe descoperirea de tipare în cadrul informaţiilor. Încă din Grecia antică, de când Hipocrate şi-a întemeiat şcoala de medicină, cu vreo 2.500 de ani în urmă, observarea, experimentarea şi analizarea informaţiilor au fost fundamentale în lupta pentru menţinerea sănătăţii.

Înainte de aceasta, medicina – în cea mai mare parte – abia dacă se deosebea de magie. Oamenii credeau că te îmbolnăveai dacă ofensai vreun zeu sau altul şi că maladia era consecinţa faptului că un spirit malefic îţi poseda

trupul. Prin urmare, activitatea unui doctor implica o mulţime de incantaţii, psalmodieri şi respectarea unor superstiţii, ceea ce sună extrem de amuzant, dar probabil că nu era deloc aşa şi pentru persoana care se baza pe toate acestea ca să nu-şi dea duhul.

Nu vreau să spun că Hipocrate ar fi vindecat lumea, de unul singur şi pentru totdeauna, de naţionalitate şi superstiţie (în definitiv, se zvonea, printre altele, că fiica sa ar fi fost un dragon de 30 de metri) 2, ci că a abordat medicina într-o manieră cu adevărat revoluţionară. În opinia lui, cauzele bolilor trebuiau înţelese prin cercetare raţională, nu prin magie. Punând accentul pe semnalarea cazurilor şi pe observaţie, el a întemeiat medicina ca ştiinţă, câştigând pe merit renumele de „părinte al medicinei moderne".3

Cu toate că explicaţiile ştiinţifice aduse de Hipocrate şi de colegii săi nu prea se mai susţin dacă le aşezăm sub lupa ştiinţei moderne (doar ei credeau că sănătatea nu era decât rezultatul armoniei, al echilibrului, dintre sânge, flegmă, bilă galbenă şi bilă neagră!) 4, concluziile la care ajungeau analizând informaţiile pe care le deţineau sunt, în mod sigur, valabile.5 (Spre exemplu, au fost primii care au tras concluzii de genul: „Pacienţii care sunt graşi de la natură sunt predispuşi să moară înaintea celor slabi") Este un laitmotiv care se regăseşte de-a lungul timpurilor. Cunoaşterea ştiinţifică e posibil s-o fi luat-o de multe ori pe căi greşite, însă progresul se naşte din capacitatea noastră de a descoperi tipare, de a clasifica simptome şi de a folosi toate aceste observaţii pentru a prezice ce poate aduce viitorul pentru un pacient.

Istoria medicinei este plină de astfel de exemple. Spre exemplu, în China secolului al XV-lea, tămăduitorii şi-au dat seama pentru prima oară că pot să imunizeze oamenii împotriva variolei. După secole de experimentare, au găsit un tipar pe care-l puteau folosi pentru a micşora de zece ori riscul de deces în urma acestei boli. Tot ce aveau de făcut era să găsească o persoană cu o formă

uşoară a bolii, să-i recolteze crustele, să le usuce, să le piseze şi să le sufle în nasul unei persoane sănătoase.6 Sau să ne gândim la epoca de aur a medicinei din secolul al XIX-lea, când aceasta adopta metode din ce în ce mai ştiinţifice, iar căutarea de tipare printre informaţii a devenit o parte integrantă a atribuţiilor unui medic. Unul dintre ei a fost maghiarul Ignaz Semmelweis, care, în anii 1840, a remarcat ceva surprinzător în statisticile privind decesele din saloanele de maternitate. Femeile care năşteau în saloane ticsite de medici erau de cinci ori mai predispuse să se îmbolnăvească de septicemie decât cele din saloanele supravegheate de moaşe. Informaţiile indicau, de asemenea, şi motivul: doctorii disecau cadavre, iar imediat după aceea se ocupau de femeile însărcinate fără să se mai oprească să se spele pe mâini! 7

Ceea ce era adevărat pentru China secolului al XV-lea şi Europa secolului al XIX-lea este adevărat şi în prezent pentru medicii din întreaga lume nu numai atunci când studiază maladiile populaţiei, ci şi când îşi îndeplinesc rolul lor principal, de zi cu zi, adică acela de a-şi îngriji pacienţii. Osul acesta este rupt sau nu? Durerea aceasta de cap este una normală sau reprezintă simptomul a ceva mai grav? Este bine să se prescrie un tratament cu antibiotice pentru a face să dispară acest abces? Toate aceste lucruri ţin de recunoaştere, clasificare şi predicţie. Aptitudini la care algoritmii se întâmplă să fie foarte, foarte buni.

Fireşte, sunt multe aspecte în profesiunea de medic pe care un algoritm nu va fi, probabil, niciodată în stare să le reproducă. Empatia, de exemplu. Sau abilitatea de a acorda sprijin pacienţilor atunci când se confruntă cu dificultăţi sociale, psihologice sau chiar financiare. Dar sunt şi domenii ale medicinei în care algoritmii pot să fie de ajutor. Mai ales în ipostazele în care recunoaşterea tiparelor medicale se manifestă în cea mai pură formă a sa, iar clasificarea şi predicţia sunt preţuite aproape până la excluderea a orice altceva. Mai ales într-un domeniu

precum patologia.

Patologii sunt medicii pe care un pacient îi întâlneşte rareori. Ori de câte ori o probă de sânge sau de ţesut este trimisă la analiză, ei sunt cei care, stând în vreun laborator îndepărtat, o vor examina şi vor completa buletinul de analiză. Contribuţia lor vine chiar la sfârşitul procesului de diagnosticare, acolo unde priceperea, precizia şi fiabilitatea sunt de o importanţă vitală. Patologii sunt adesea cei care îţi spun dacă ai sau nu cancer. Prin urmare, dacă biopsia pe care o analizează este singurul lucru care te desparte de chimioterapie, intervenţie chirurgicală sau chiar ceva mai rău, vrei să fii sigur că ei fac lucrurile cum trebuie.

Iar munca lor nu este deloc uşoară. Potrivit sarcinilor care îi revin, un patolog oarecare va examina sute de lamele pe zi, fiecare conţinând zeci de mii – uneori chiar sute de mii – de celule care stau în suspensie între plăci de sticlă micuţe. Este cel mai dificil joc de tip *Unde este Wally*?[38] pe care ţi-l poţi imagina. Patologii trebuie să examineze cu meticulozitate fiecare probă în căutarea unor anomalii infime care s-ar putea ascunde oriunde în imensa galaxie de celule pe care o văd ei sub lentilele microscopului.

„Este o sarcină extraordinar de dificilă”, a declarat Andy Beck\ patolog la Universitatea Harvard şi fondatorul PathAI, o companie înfiinţată în 2016 care creează algoritmi pentru clasificarea lamelelor de biopsie. „Dacă fiecare patolog ar trebui să examineze cu foarte mare atenţie cinci lamele pe zi, ai putea să crezi că ar reuşi să atingă perfecţiunea. Dar nu aceasta este lumea în care trăim”.

În mod cert, nu este. Iar în lumea reală, sarcina

38 Colecţie de cărţi pentru copii creată de ilustratorul britanic Martin Handford, în care personajul numit Wally se ascunde în desenele ce înfăţişează diferite mulţimi, iar cititorii trebuie să-l descopere. Editorii din diverse ţări i-au dat personajului nume diferite, astfel încât acesta este cunoscut drept Waldo în SUA şi Canada, drept Charlie, în Franţa, Walter în Germania, Aii în Turcia, Efi în Israel şi Willy în Norvegia (n. tr.).

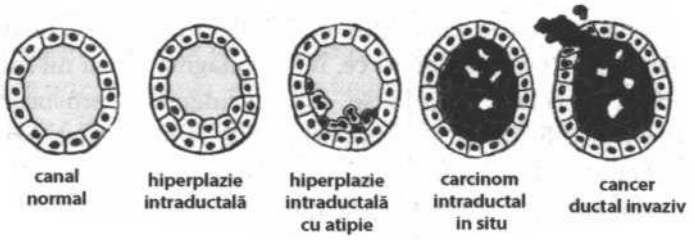
patologilor devine şi mai dificilă din cauza complexităţii irustrante a biologiei. Să ne întoarcem la exemplul cu porumbeii şi cancerul mamar pe care aceştia reuşeau atât de bine să îl detecteze. A decide dacă o persoană suferă de această afecţiune nu înseamnă să dai un răspuns simplu, precum „da” sau „nu”, deoarece diagnosticele de cancer mamar acoperă un întreg spectru. La unul din capete se află probele benigne, în care celulele normale arată exact aşa cum ar trebui să arate. La capătul opus se află tumorile de tipul cel mai urât – carcinoamele invazive, în care celulele canceroase au migrat din canalele mamare şi au început să prolifereze în ţesutul înconjurător. Cazurile situate la aceste extreme sunt relativ uşor de depistat. Un studiu recent arăta că patologii reuşesc să diagnosticheze corect 96% din specimenele care sunt, în mod clar, maligne, o precizie aproximativ echivalentă cu aceea pe care a atins-o şi grupul de porumbei când i s-a prezentat o sarcină similară.9

Între aceste două extreme, însă – între absolut normal şi malign într-un mod îngrozitor de clar –, se situează alte câteva categorii, mai ambigue, aşa cum se poate vedea în diagrama prezentată în continuare. În proba dumneavoastră s-ar putea să existe un grup de celule atipice care arată cam dubios, dar care nu înseamnă neapărat ceva îngrijorător. Ar putea fi nişte formaţiuni precanceroase care s-ar putea transforma (sau nu!) în ceva grav. Sau s-ar putea să existe un cancer care încă nu s-a răspândit dincolo de canalele mamare (aşa numitul carcinom *in situ*[39]).

Cărei categorii anume se va considera că aparţine proba dumneavoastră va avea probabil un impact imens asupra tratamentului pe care îl veţi urma. În funcţie de locul din spectru în care se situează proba, medicul v-ar

39 Carcinomul in situ (CIS) este o formă timpurie de cancer, definită prin absenţa invaziei de celule tumorale în ţesutul înconjurător. Cu alte cuvinte, celulele neoplazice proliferează în habitatul lor normal, de unde şi denumirea de in situ (lat.), „în locul său”, putând fi extirpate complet (n. tr.).

putea sugera orice, de la mastectomie până la niciun fel de intervenţie.



Problema este că poate fi extrem de dificil să faci deosebirea între aceste categorii ambigue. Până şi patologii cu experienţă pot avea opinii diferite în privinţa diagnosticării corecte a unei anumite probe. Pentru a verifica în ce măsură variază opiniile medicale, un studiu din 2015 a selectat 72 de biopsii ale ţesutului mamar despre care se considera că ar conţine celule cu anomalii benigne (o categorie situată pe la mijlocul spectrului) şi apoi s-au solicitat părerile unui număr de 115 patologi. În mod îngrijorător, doar în 48% dintre cazuri patologii au emis diagnostice similare.10

Şi, de vreme ce, pare-se, şansele de diagnosticare corectă sunt de jumi-juma, am putea la fel de bine să dăm cu banul pentru a stabili diagnosticul. Cap, şi te-ai putea alege cu o mastectomie inutilă (care să te coste sute de mii de dolari dacă trăieşti în SUA). Pajură, şi ai putea să ratezi şansa de a prinde cancerul în stadiul incipient. Oricum ar fi, impactul poate fi devastator.

Evident, când miza este atât de mare, cel mai mult contează acurateţea. Aşadar, un algoritm poate să obţină rezultate mai bune decât acestea?

MAŞINI CARE VĂD

Până de curând, construirea unui algoritm care să poată recunoaşte ceva, indiferent ce, într-o imagine, ca să nu mai vorbim de celule canceroase, era considerată o provocare extrem de dificilă. Nu avea importanţă că oamenii înţelegeau fotografiile cu atâta uşurinţă; explicarea exactă a *modului în care reuşeau* să facă acest lucru s-a dovedit o sarcină neînchipuit de dificilă.

Ca să înţelegeţi de ce, imaginaţi-vă că scrieţi instrucţiuni pentru a-i comunica unui computer dacă într-o fotografie apare sau nu un câine. Aţi putea începe cu ceea ce este evident: dacă are patru picioare, urechi pleoştite, blană şi aşa mai departe. Dar ce ar trebui făcut cu acele fotografii în care câinele stă jos? Sau cu acelea în care nu i se văd toate picioarele? Dar cu pozele care înfăţişează un câine cu urechi ascuţite? Sau unul cu urechile pleoştite doar pe jumătate? Sau altul care nu priveşte deloc spre obiectivul aparatului de fotografiat? Şi prin ce diferă „blana” unui câine dintr-o fotografie de imaginea unui covor cu fire lung? Sau de lâna unei oi? Sau de iarbă?

Sigur că aţi putea introduce toate aceste detalii ca instrucţiuni suplimentare şi să faceţi apoi algoritmul să treacă prin fiecare tip posibil de urechi de câine, blană de câine sau poziţie şezândă, numai că, în scurt timp, algoritmul va deveni atât de voluminos, încât va fi complet nefuncţional înainte chiar ca dumneavoastră să fi început măcar să faceţi deosebirea între câini şi alte creaturi cu blană şi patru picioare. Aşadar, trebuie să găsiţi o altă modalitate de diferenţiere. Secretul este să renunţaţi la o paradigmă bazată pe reguli şi să folosiţi ceva ce se numeşte o „reţea neurală”.11

Vă puteţi imagina reţeaua neurală ca pe o imensă structură matematică în care apar o mulţime de butoane şi cadrane. La un capăt se introduce poza, aceasta parcurge întreaga structură, iar la capătul celălalt iese o presupunere despre ce ar putea conţine imaginea respectivă. O probabilitate pentru fiecare categorie: Câine sau Non-Câine.

La început, reţeaua neurală va fi o mare harababură. Începe prin a nu şti nimic – nu are habar ce este sau ce nu este un câine. Toate butoanele şi cadranele sunt setate în poziţia „aleatoriu”. În consecinţă, răspunsurile pe care le va oferi vor fi complet aiurea – reţeaua nu ar putea recunoaşte o imagine nici dacă însăşi sursa ei de alimentare cu electricitate ar depinde de asta. Însă, cu

fiecare fotografie pe care o veţi introduce, veţi ajusta butoanele şi cadranele acelea. Treptat, o veţi antrena.

Şi acum puteţi să introduceţi imaginea unui câine. După fiecare supoziţie pe care o va face reţeaua, un set de reguli matematice va intra în funcţiune pentru a ajusta toate butoanele, până când predicţia se va apropia de răspunsul corect. Atunci veţi introduce o altă imagine şi apoi încă una, făcând ajustări de fiecare dată când reţeaua neurală va înţelege ceva greşit, întărind căile prin seria de butoane care conduc la succes şi reducând, treptat, importanţa acelora care au condus la eşec. Informaţia despre ceea ce face ca o poză reprezentând un câine să semene cu o altă poză înfăţişând un câine se propagă în sens invers prin reţea. Procesul continuă până când – după introducerea a sute şi mii de fotografii – se primesc tot mai puţine răspunsuri greşite cu putinţă. În cele din urmă, îi puteţi arăta reţelei o imagine pe care nu a văzut-o niciodată şi aceasta va reuşi să afirme, cu un nivel ridicat de acurateţe, dacă aceasta înfăţişează sau nu un câine.

Surprinzător în ceea ce priveşte reţelele neurale este faptul că utilizatorii acestora, de obicei, nu înţeleg cum sau de ce ajunge algoritmul la concluziile sale. O reţea neurală care clasifică imagini de câini nu funcţionează prin identificarea trăsăturilor pe care dumneavoastră sau eu le-am putea recunoaşte ca fiind potenţial canine. Ea nu caută o unitate de măsură pentru „*chihuahuaism*” sau „marele danezism” –, totul este mult mai abstract decât atât: reţeaua neurală detectează tipare ale marginilor şi ale luminii şi umbrei din fotografii, ceea ce nu este foarte inteligibil pentru un observator uman (aruncaţi o privire asupra exemplului de recunoaştere a imaginii din capitolul „Puterea” ca să înţelegeţi mai bine ce vreau să spun). Pentru că unei fiinţe umane îi este greu să conceptualizeze acest proces, programatorii nu ştiu decât că au setat algoritmul să dea răspunsurile corecte; ei nu ştiu neapărat în amănunt cum ajunge acesta la rezultatul respectiv.

Iată încă un „algoritm de învăţare automată”, ca

acele „păduri aleatorii” (*random forests*) pe care le-am întâlnit la capitolul „Justiţia”. El merge dincolo de ceea ce inginerul îl programează să execute şi învaţă singur din imaginile care i se furnizează. Această capacitate de învăţare înzestrează algoritmul cu o remarcabilă „inteligenţă artificială”. Iar multiplele niveluri de butoane şi cadrane îi conferă reţelei o structură în profunzime, de unde şi sintagma *deep leaming*[40].

Reţelele neurale există de pe la mijlocul secolului al XX-lea, însă până de curând ne-a lipsit accesul pe scară largă la calculatoarele cu adevărat puternice necesare pentru a obţine tot ce este mai bun de la aceste reţele. Lumea a fost, în cele din urmă, nevoită să îşi manifeste poziţia şi să le ia în serios în 2012, când cercetătorul în calculatoare Geoffrey Hinton şi doi studenţi de-ai să-i au înscris un nou tip de reţea neurală într-un concurs de recunoaştere a fotografiilor.12 Provocarea care li se lansase – printre altele – era recunoaşterea câinilor. Algoritmul lor înzestrat cu inteligenţă artificială i-a făcut praf până şi pe cei mai buni concurenţi şi a declanşat un reviriment de mari proporţii în domeniul *deep leaming*.

Un algoritm care lucrează fără ca noi să ştim cum ia deciziile s-ar putea să semene a scamatorie, dar poate că acest lucru nu este chiar atât de diferit de modul în care învăţăm noi înşine. Gândiţi-vă la comparaţia următoare. O echipă a instruit recent un algoritm să facă deosebirea între fotografiile cu lupi şi acelea în care apar câini husky. Apoi au arătat cum, datorită modului în care îşi reglase propriile cadrane, algoritmul nu folosea absolut deloc indicii care să aibă legătură cu vreun câine. Îşi baza răspunsul pe informaţia dacă imaginea avea sau nu zăpadă în fundal. Plus zăpadă = lup. Fără zăpadă = husky.13

La puţin timp după publicarea lucrării lor, stăteam de vorbă cu Frank Kelly, un profesor de matematică de la Universitatea Cambridge, care mi-a relatat o conversaţie pe care o avusese cu nepotul său. Îşi ducea nepotul de

40 „învăţare profundă” (engl.) (vezi glosarul) (n. tr.).

patru ani la grădiniţă, când au trecut pe lângă un câine husky. Nepotul său a remarcat că acel câine „arăta” ca un lup. Când Frank l-a întrebat de unde ştia că nu era un lup, micuţul îi răspunsese: „Pentru că e în lesă”.

Sunt două lucruri de dorit de la un algoritm performant de scanare a cancerului mamar. În primul rând, să fie suficient de *sensibil* ca să detecteze anomaliile prezente la toţi sânii care prezintă tumori, fără să-i scape vreun pixel din imagine şi fără să anunţe apoi că nu există anomalii. Dar, în al doilea rând, mai trebuie şi ca algoritmul să fie destul de *precis*, să dea dovadă de *specificitatef* astfel încât să nu semnaleze că o porţiune de ţesut mamar perfect normal ar fi demn să stârnească suspiciuni.

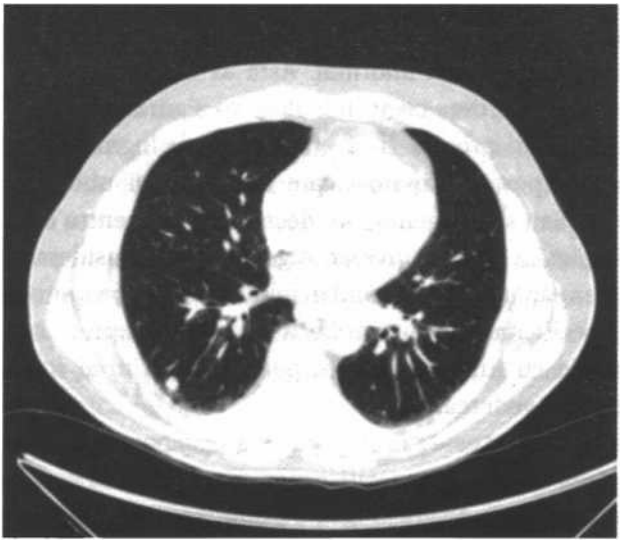
Am întâlnit deja principiile sensibilităţii şi specificităţii în capitolul „Justiţia”. Sunt rude apropiate cu *fals negativele* şi *fals pozitivele* (sau cu Darth Vader-ii şi a Luke Skywalker-ii – denumirile care, dacă mă întrebaţi pe mine, ar trebui să fie folosite, în mod oficial, şi în literatura ştiinţifică). În contextul nostru, al diagnosticării medicale, un fals pozitiv apare atunci când unei femei sănătoase i se spune că are cancer la sân, iar un fals negativ – atunci când o femeie cu tumori primeşte diagnosticul că ar fi sănătoasă. Un test specific aproape că nu va avea rezultate fals pozitive, pe când unul sensibil va avea puţine rezultate fals negative. Nu are importanţă în ce context lucrează algoritmul – dacă prognozează că un anumit infractor va recidiva sau nu, dacă diagnostichează cancerul mamar sau (cum vom vedea în capitolul „Infracţionalitatea”) dacă identifică tiparele activităţii infracţionale – povestea este mereu aceeaşi. E de dorit să avem cât mai puţine rezultate fals pozitive şi fals negative cu putinţă.

Problema este că perfecţionarea unui algoritm înseamnă adesea o alegere între sensibilitate şi specificitate. Dacă vă concentraţi asupra îmbunătăţirii uneia, asta înseamnă adesea o scădere a performanţelor celeilalte. Dacă, de exemplu, aţi decis să daţi întâietate

eliminării complete a fals negativelor, algoritmul ar putea să semnaleze fiecare ţesut mamar pe care îl vede drept anormal. Asta ar însemna că are 100% sensibilitate, ceea ce ar îndeplini, cu siguranţă, obiectivul pe care vi l-aţi propus. Însă ar însemna şi îngrozitor de mulţi oameni perfect sănătoşi care ar fi supuşi unui tratament inutil. Sau să spunem că aţi decis să optaţi pentru eliminarea completă a fals pozitivelor. Algoritmul ar considera pe toată lumea sănătoasă, atingând în felul acesta un procent de 100% la specificitate. Minunat! Dacă nu cumva sunteţi una dintre femeile cu tumori mamare pe care algoritmul tocmai le-a trecut cu vederea.

Interesant este că patologii umani nu par să aibă probleme în ceea ce priveşte specificitatea. Ei aproape că nu identifică niciodată în mod greşit celule sănătoase ca fiind canceroase, însă oamenii au o oarecare problemă cu sensibilitatea: pur şi simplu, se pare că ne este îngrijorător de uşor să trecem cu vederea micile tumori, ba chiar şi pe cele care sunt, în mod evident, maligne.

Aceste slăbiciuni omeneşti au fost scoase la iveală într-o competiţie desfăşurată recent în care fiinţele umane s-au confruntat cu algoritmii. Echipe de programatori din toată lumea s-au luat la întrecere cu patologi, să găsească toate tumorile existente pe patru sute de lamele, într-un concurs cunoscut sub numele de CAMELYON16. Ca să uşureze lucrurile, toate cazurile se situau la una dintre cele două extreme: ţesut perfect normal sau cancer mamar invaziv, iar patologii nu aveau nicio limită temporală: dispuneau de tot timpul de care aveau nevoie pentru a examina biopsiile. Aşa cum era de aşteptat, oamenii au pus diagnosticul general în mod corect (cu o precizie de 96%) 14 şi fără să identifice niciun fals pozitiv. Însă le-au scăpat şi o mulţime de mici celule canceroase ascunse prin ţesut, reuşind să observe numai 73% dintre acestea în 30 de ore de examinare.



Problema nu a fost neapărat numărul efectiv de pixeli care trebuiau verificaţi. Oamenii pot cu uşurinţă să nu observe anomalii foarte evidente chiar şi atunci când se uită direct la ele. În anul 2013, cercetătorii de la Harvard au ascuns în secret imaginea unei gorile într-o serie de radiografii toracice şi au cerut unui număr de 24 de radiologi, care nu bănuiau nimic, să verifice radiogramele respective pentru a depista semne de cancer. Ei bine, 83% dintre aceştia nu au observat gorila, în pofida faptului că sistemul de urmărire a privirii folosit de oamenii de ştiinţă a arătat că majoritatea se uitaseră chiar la ea.15 Puteţi încerca şi dumneavoastră să o depistaţi, în poza de mai sus.16

Algoritmii au exact problema opusă: vor marca imediat grupurile de celule cu anomalii, ba chiar şi pe unele perfect sănătoase. În timpul studiului CAMELYON16, de exemplu, cea mai performantă reţea neurală participantă a reuşit să identifice un procent impresionant (92, 4%) dintre tumori17, însă a făcut astfel opt erori de identificare fals pozitivă per lamelă, prin semnalarea greşită a unor grupuri de celule normale ca fiind canceroase. Cu o specificitate atât de scăzută, algoritmii actuali care reprezintă cel mai avansat stadiu al

cercetării înclină, în mod cert, în diagnosticare, spre abordarea „toată lumea are cancer mamarf, ceea ce face, pentru moment, ca ei să nu fie, practic, suficient de performanţi pentru a putea să emită propriile rapoarte patologice.

Veştile bune sunt, însă, că nici nu ne propunem să le cerem acest lucru! Intenţionăm să combinăm, în schimb, calităţile omului cu acelea ale maşinii. Algoritmul efectuează munca de rutină – examinarea volumului imens de informaţii de pe lamele – evidenţiind câteva arii principale de interes. Apoi patologul preia comanda. Nu are importanţă dacă maşina semnalează celule care nu sunt canceroase, pentru că specialistul uman va face o verificare rapidă, eliminând tot ce este normal. Genul acesta de parteneriat în care algoritmul realizează o preexaminare, o presortare a informaţiilor, nu numai că economiseşte foarte mult timp, ci creşte şi precizia generală a diagnosticului până la un procent uimitor de 99, 5%.18

Oricât de miraculos vi s-ar părea, adevărul este că patologii umani au fost întotdeauna buni la diagnosticarea tumorilor canceroase agresive. Cazurile dificile pentru ei sunt cele ambigue, de la mijlocul spectrului, în care deosebirea dintre celulele canceroase şi cele sănătoase este mai subtilă. Pot, oare, algoritmii să ne fie de ajutor şi în acest caz? Răspunsul este (probabil) da. Însă nu încercând să stabilească un diagnostic prin utilizarea categoriilor complicate pe care patologii le-au folosit dintotdeauna. Poate că algoritmul – care este cu mult mai performant în depistarea anomaliilor ascunse în mici fragmente de informaţie – ar putea să ofere, în schimb, o cu totul altă modalitate, mai bună, de diagnosticare. Făcând ceva ce medicii umani nu pot.

STUDIUL REALIZAT PE CĂLUGĂRIŢE[41]

41 Nun Study of Aging and Alzheimer’s Disease (engl.), „Studiul privind îmbătrânirea şi [instalarea maladiei lui] Alzheimer realizat pe călugăriţe”, demarat în 1986 de David Snowdon (n. red.).

În anul 1986, David Snowdon, un epidemiolog de la Universitatea din Kentucky, a reuşit să convingă 678 de călugăriţe să-i cedeze creierele după moarte. Măicuţele, care aparţineau Ordinului Surorilor de Notre Dame, au acceptat să participe la o cercetare ştiinţifică ieşită din comun despre cauzele apariţiei maladiei Alzheimer.

Fiecare dintre aceste femei, care la începutul studiului aveau vârste cuprinse între 75 şi 103 ani, avea să efectueze o serie de teste de memorie în fiecare an pe durata întregii vieţi. Apoi, după moarte, creierele lor aveau să fie donate proiectului respectiv. Indiferent dacă manifestau sau nu simptome de demenţă, ele au promis că vor permite echipei lui Snowdon să înlăture cel mai preţios organ al lor pentru a fi analizat în căutarea semnelor acestei afecţiuni.19

Generozitatea călugăriţelor a dus la crearea unui remarcabil set de date. Din moment ce niciuna dintre ele nu avea copii, nu fuma şi nu bea prea mult, oamenii de ştiinţă au putut să excludă mulţi dintre factorii externi despre care se credea că ar fi dus la creşterea probabilităţii de a contracta maladia Alzheimer. Şi, din moment ce aveau un stil de viaţă asemănător, cu acces egal la sistemul de sănătate şi la cel de asistenţă socială, călugăriţele ofereau ele însele şi grupul de control.

Jocul a decurs bine când, după câţiva ani de la începerea studiului, echipa a descoperit că acest grup experimental mai oferea un tezaur de informaţii la care putea avea acces. În tinereţe, multora dintre măicuţele acum în etate li se ceruse, înainte de a li se permite să se călugărească, să înmâneze Ordinului un eseu autobiografic scris de mână. Aceste eseuri fuseseră scrise când femeile aveau – în medie – doar 22 de ani, cu zeci de ani înainte ca oricare dintre ele să manifeste vreun simptom de demenţă. Cu toate acestea, în mod uimitor, oamenii de ştiinţă au descoperit în scrisul lor de mână indicii care preziceau ce avea să li se întâmple într-un viitor îndepărtat.

Cercetătorii au analizat complexitatea limbajului din

fiecare eseu şi au găsit o legătură între cât de elocvente erau călugăriţele ca tinere femei şi şansele ca ele să dezvolte demenţă la o vârstă înaintată.

Iată, de exemplu, un citat dintr-un eseu al unei călugăriţe care şi-a păstrat abilităţile cognitive excelente întreaga sa viaţă:

După ce am terminat clasa a VIII-a, în 1921, am dorit să devin aspirantă la Mankato, însă nu am avut curajul să cer eu însămi permisiunea părinţilor, aşa încât sora Agreda a făcut-o în locul meu, iar aceştia şi-au dat imediat consimţământul.

Comparaţi această frază cu următoarea, scrisă de o altă călugăriţă, ale cărei punctaje de acuitate a memoriei au decăzut constant în ultimii ani de viaţă:

După ce am absolvit şcoala, am lucrat la oficiul de poştă.

Asocierea era atât de clară, încât oamenii de ştiinţă puteau de-acum să prezică de-a dreptul care dintre călugăriţe ar fi putut avea demenţă numai citindu-le scrisorile. 90% dintre cele care au dezvoltat mai târziu maladia Alzheimer aveau o „capacitate lingvistică limitată” ca tinere femei, pe când numai 13% dintre călugăriţele care şi-au păstrat capacităţile lingvistice până la o vârstă înaintată au primit un punctaj de „densitate scăzută a ideilor” la eseurile lor.20

Unul dintre lucrurile pe care acest studiu le pune în lumină este cantitatea incredibilă de informaţii pe care încă o avem de aflat despre corpurile noastre. Chiar dacă ştim că acea legătură ar putea să existe, nu ştim şi *de ce*. (Oare o bună educaţie este suficientă pentru a împiedica apariţia demenţei? Sau persoanele predispuse să se îmbolnăvească de Alzheimer se simt mai în largul lor folosind un limbaj mai simplu?) Dar asta ar sugera că maladia Alzheimer se dezvoltă pe o durată de câţiva zeci de ani!

Încă şi mai important pentru ceea ce ne propunem, asta demonstrează că indicii subtile despre viitoarea

noastră stare de sănătate se pot ascunde chiar şi în cele mai mărunte, mai neaşteptate fragmente de informaţie, cu ani şi ani înainte ca noi să manifestăm simptomele vreunei maladii. Aceasta ne oferă o idee despre cât de puternici ar putea fi algoritmii medicali care vor fi folosiţi în viitor pentru analizarea informaţiilor. Poate că, într-o bună zi, vor reuşi să detecteze semnele cancerului cu câţiva ani înaintea medicilor.

PUTEREA PREDICŢIEI

La sfârşitul anilor 1970, un grup de patologi din Ribe, un oraş din Danemarca, a început să efectueze duble mastectomii pe un lot de cadavre. Femeile decedate aveau vârste cuprinse între 22 şi 89 de ani, iar şase din totalul de 83 muriseră de cancer mamar invaziv. Bineînţeles că atunci când cercetătorii au pregătit sânii extirpaţi în vederea examinării lor de către un patolog – împărţind fiecare sân în patru bucăţi, apoi tăind ţesutul în felii subţiri pentru a-l aşeza pe lamele –, cele şase eşantioane au manifestat semnele maladiei. Dar, spre uimirea cercetătorilor, din restul de 77 de femei – care muriseră din alte cauze, fără nicio legătură cu cancerul, printre care boli de inimă şi accidente de automobil –, aproape un sfert prezentau semnele de avertizare ale cancerului mamar pe care patologii le căutau la pacientele în viaţă.

Fără să manifeste niciun semn de boală, 14 femei aveau celule canceroase *in situ* care nu se dezvoltaseră niciodată dincolo de canalele sau glandele mamare. Celule care ar fi fost considerate drept cancer mamar malign dacă femeile ar fi fost în viaţă. Trei aveau celule atipice care ar fi fost, de asemenea, considerate a reprezenta un motiv de îngrijorare la o biopsie, iar o femeie suferea chiar de cancer mamar invaziv, deşi nu avusese nici cea mai vagă idee despre acest lucru atunci când murise.21

Cifrele au fost surprinzătoare, dar studiul nu era doar o coincidenţă. Şi alţi cercetători găsiseră rezultate asemănătoare. De fapt, unii chiar estimau că, în orice moment, în jur de 9% dintre femei ar putea să trăiască fără

să ştie că au dezvoltat tumori mamare22 – cam de zece ori mai mult faţă de cele care sunt diagnosticate în prezent cu cancer mamar.23

Aşadar, ce se întâmplă? Avem de-a face cu o epidemie tăcută? Conform doctorului Jonathan Kanevsky, un inovator în medicină şi chirurg rezident la Universitatea McGill din Montreal, răspunsul este: „Nu. Nu chiar”. Pentru că prezenţa cancerului nu este neapărat o problemă:

Dacă are cineva o celulă canceroasă în corp, sunt şanse ca sistemul imunitar să o identifice drept o celulă mutantă, să o atace şi să o distrugă, iar acel cancer nu se va dezvolta niciodată, devenind ceva grav. Uneori, însă, sistemul imunitar se zăpăceşte, adică organismul favorizează dezvoltarea cancerului, permiţându-i să prolifereze. În acest moment, cancerul poate să ucidă.24

Nu toate tumorile sunt la fel. De unele se va ocupa organismul, altele vor rămâne liniştite acolo până la trecerea în nefiinţă, iar altele ar putea să devină un cancer agresiv în toată regula. Problema este că, adesea, avem foarte puţine căi de a şti ce se va întâmpla în fiecare caz în parte.

Este şi motivul pentru care aceste categorii dificile, situate între benign şi malign pot fi o problemă. Sunt singurele clasificări pe care medicii le au la dispoziţie, dar, dacă unul dintre ei găseşte în biopsia dumneavoastră un grup de celule care îi dau, cât de cât, de bănuit, eticheta pe care va alege să i-o pună poate fi utilă doar pentru a descrie ce există în ţesut în prezent. Asta nu este de prea mare ajutor atunci când trebuie să-ţi furnizeze indicii despre viitor. Şi fireşte că viitorul îi preocupă cel mai mult pe pacienţii îngrijoraţi.

În consecinţă, oamenii sunt, adesea, extrem de precauţi în privinţa tratamentului pe care îl aleg. Să luăm, de exemplu, cancerele *in situ*. Această categorie este plasată spre capătul mai îngrijorător al spectrului, căci deja este prezentă o excrescenţă canceroasă care încă nu

s-a răspândit în ţesutul înconjurător. Oricât de grav ar suna prognosticul de „cancer *in situ*”, numai aproximativ unul din zece se va transforma în ceva care ar putea să ucidă. Nu este mai puţin adevărat că, în Statele Unite, un sfert dintre femeile care primesc acest diagnostic vor fi supuse unei mastectomii totale – o intervenţie majoră care le schimbă viaţa nu doar în plan fizic, ci, adesea, şi în cel emoţional.25

Din ceea ce s-a constatat, cu cât faci examinări mai insistente pentru depistarea cancerului mamar, cu atât afectezi mai multe femei care şi-ar fi trăit mai departe fericite viaţa fără să aibă habar de existenţa unor tumori inofensive. Un grup independent din Marea Britanie a ajuns la concluzia că, pentru fiecare 10.000 de femei care vor fi chemate să-şi facă un examen mamografic în următorii 20 de ani, vor fi prevenite 43 de decese de cancer mamar. Iar un studiu publicat în *New England Journal of Medicine* a ajuns la concluzia că, pentru fiecare 10.000 de femei care efectuează mamografii de rutină, unui număr de doar 30 le vor fi depistate tumori care ar putea să le pună în pericol viaţa.26 Numai că – depinzând de statisticile la care te raportezi – de trei sau patru ori mai multe femei vor fi supradiagnosticate, fiind tratate pentru tumori care nu ar fi devenit niciodată letale.27

Problema supradiagnosticării şi a tratării excesive este dificil de rezolvat atunci când te pricepi la depistarea anomaliilor, dar nu şi la prezicerea modului în care vor evolua ele. Chiar şi aşa, ar putea exista speranţe. Poate că – la fel ca în cazul călugăriţelor care scriseseră eseuri în tinereţe – mici indicii despre modul în care va evolua sănătatea unei persoane în viitor, peste ceva ani, se pot ascunde şi pot fi depistate în informaţii trecute şi prezente despre aceasta. Astfel, extragerea acestor informaţii ar fi sarcina perfectă pentru o reţea neurală.

Într-un domeniu în care medicii s-au chinuit zeci de ani să descopere *de ce* o anumită anomalie este mai periculoasă decât alta, un algoritm care să nu fie instruit

să caute *ceva anume* ar putea să-şi arate adevărata valoare. Atât timp cât poţi strânge un număr foarte mare de lamele de biopsie (care să includă câteva probe de tumori care au format metastaze – adică s-au extins la alte organe din corp – şi, de asemenea, câteva care nu au format metastaze) pentru antrenarea reţelei tale neurale, aceasta ar putea căuta indicii ascunse cu privire la starea de sănătate viitoare, complet liberă de oricare dintre prejudecăţile care vin la pachet cu teoria. Jonathan Kanevsky considera:

Ar depinde de algoritm să stabilească caracteristicile particulare din fiecare imagine care prezic dacă o tumoare va evolua sau nu spre metastază.28

Dacă ar exista un astfel de algoritm, categoria în care se încadrează biopsia dumneavoastră ar deveni mult mai puţin importantă. Nu v-ar mai preocupa „de ce" şi „cum", aţi putea sări direct la informaţia care contează: aveţi sau nu nevoie de tratament?

Vestea cea bună este că munca pentru crearea unui asemenea algoritm a început deja. Andy Beck, patologul de la Harvard şi preşedintele companiei PathAI, pe care am menţionat-o mai devreme, şi-a instruit algoritmul să funcţioneze pe o serie de probe ale unor pacienţi din Olanda şi a descoperit că, de fapt, cei mai buni indicatori capabili să prezică şansele de supravieţuire nu se găseau în cancerul ca atare, ci în anomaliile din ţesuturile învecinate zonei canceroase.29 Este o descoperire importantă – un exemplu concret al modului în care algoritmii înşişi duc cercetarea mai departe, dovedind că pot să descopere tipare care să ne îmbunătăţească puterea de predicţie.

Dispunem, în prezent, de o bogăţie incredibilă de informaţii pe care le putem valorifica. Datorită mamografiilor de rutină care se efectuează în întreaga lume avem, probabil, mai multe imagini ale ţesutului mamar decât ale oricărui alt organ din corp. Eu nu sunt patolog, însă fiecare specialist cu care am discutat m-a

convins că suntem foarte aproape de a reuşi să prezicem cu toată încrederea dacă un eşantion problematic va deveni sau nu canceros. Sunt chiar şanse mari ca, până la lansarea în librării a acestei cărţi, cineva, undeva, să fi transformat deja această idee revoluţionară în realitate.

DIAGNOSTICAREA DIGITALĂ

Aceste idei au aplicaţii care depăşesc diagnosticarea cancerului mamar. Reţelelor neurale pe care le construiesc Andy Beck şi alţii nu le pasă, în mod deosebit, ce anume examinează: câini, pălării sau bucăţi de brânză. Atât timp cât le anunţi când au dreptate şi când greşesc, ele vor învăţa. Iar acum, când această familie de algoritmi este suficient de performanţă pentru a fi utilizată, ea are impact într-o mulţime de domenii ale medicinei moderne.

Un succes deosebit de important a dobândit de curând echipa Google Brain, de exemplu, care a construit un algoritm capabil să depisteze principala cauză a orbirii la nivel mondial, retinopatia diabetică, o afecţiune care poate fi prevenită. Această maladie afectează vasele de sânge din zonele sensibile la lumină ale ochiului. Dacă ştii că suferi de aşa ceva, ţi se pot prescrie injecţii pentru salvarea vederii, însă, dacă nu este depistată la timp, retinopatia diabetică poate duce la orbire ireversibilă. În India, unde accesul la specialişti care să diagnosticheze această afecţiune este limitat, 45% dintre persoanele cu retinopatie diabetică îşi vor pierde, într-o anumită măsură, vederea chiar înainte de a afla că suferă de aşa ceva. Algoritmul echipei Google, care a fost construit în colaborare cu medici din India, este în prezent tot atât de bun pentru diagnosticarea acestei afecţiuni ca un medic oftalmolog.

În mod similar, există algoritmi care depistează afecţiunile cardiovasculare30, emfizemul pulmonar31, accidentele vasculare cerebrale32 şi melanoamele pielii33. Există chiar şi sisteme care diagnostichează în timp real[42]

42 Real time (engl.), atunci când este vorba despre un sistem de prelucrare a datelor, această expresie semnifică faptul că rezultatele

polipii în timpul unei colonoscopii pe viu.

Adevărul este că, dacă poţi să fotografiezi ceva şi să-i pui o etichetă, atunci poţi crea şi un algoritm care să descopere acel lucru. Şi este posibil să obţii un diagnostic mai precis (şi care să fie depistat, poate, în fazele incipiente ale bolii) decât ar putea s-o facă orice medic uman.

Ce se întâmplă, însă, cu informaţiile medicale care apar într-o formă mai dezordonată? Succesul acestor algoritmi de diagnosticare poate fi dus mai departe, dincolo de acest gen de sarcini specializate şi concentrate asupra unor sarcini restrânse? Ar putea, oare, o maşină să atribuie un înţeles notiţelor mâzgălite de un medic, de exemplu? Sau să sesizeze indiciile mărunte din felul în care un pacient îşi descrie durerea pe care o resimte?

Ce spuneţi, însă, de cea mai nouă fantasmagorie ştiinţifico-fantastică din domeniul sănătăţii, în care o maşină aflată în cabinetul medicului dumneavoastră de familie ascultă cu atenţie simptomele pe care le acuzaţi şi vă analizează antecedentele medicale? Să îndrăznim să ne imaginăm un computer care stăpâneşte fiecare fragment de cercetare medicală de ultimă oră? Unul care oferă un diagnostic exact şi un plan de tratament perfect adaptat?

Pe scurt: ce-aţi spune de ceva care să semene cu Watson de la IBM?

În anul 2004, Charles Lickel se înfrupta dintr-o friptură împreună cu alţi colegi de-ai să-i ia un restaurant din New York. Pe la jumătatea mesei, zona aceea a localului a început să se golească. Intrigat, Charles a urmărit mesenii şi a descoperit că se înghesuiseră în jurul unui televizor, urmărind cu interes popularul concurs american *Jeopardy*. Celebrul câştigător al concursului, Ken Jennings, avea şansa să continue seria celor şase luni de victorie neîntreruptă care depăşea toate recordurile, iar clienţii localului nu doreau să rateze această

---

sunt obţinute concomitent cu evenimentul care a generat datele de intrare (inputul) (n. red.).

performanţă.34

Charles Lickel era vicepreşedintele responsabil de software al IBM. în ultimii câţiva ani, de când Deep Blue îl învinsese pe Garry Kasparov la şah, şefii de la IBM îl pisau pe Charles să găsească o nouă provocare demnă de atenţia companiei. Pe când se afla în restaurantul din New York, urmărindu-i pe clienţii captivaţi de acel campion uman al concursului *Jeopardy*, Lickel a început să se întrebe dacă nu putea fi construit un calculator care să-l învingă.

Nu avea să fie uşor. Construirea maşinii cunoscute drept „Watson”, pe care şi-a imaginat-o Charles în acel restaurant, a durat şapte ani. Dar, în cele din urmă, Watson avea să-l provoace la o competiţie pe Ken Jennings, într-un episod special din *Jeopardy* şi să-l învingă fără echivoc la jocul care îi adusese notorietatea. Între timp, IBM a început construirea primului calculator de diagnosticare atotştiutor din lume, la care ne vom întoarce în curând. Dar, mai întâi, permiteţi-mi să vă prezint câteva aspecte esenţiale despre maşina care a câştigat concursul *Jeopardy*, care au stat şi la baza algoritmilor de diagnosticare medicală.

Pentru cei care nu au auzit niciodată de ea, *Jeopardy* este o foarte cunoscută emisiune de divertisment americană sub forma unui concurs de întrebări inversat: concurenţilor li se oferă indicii sub formă de răspunsuri, iar ei trebuie să formuleze întrebarea care a stat la baza acestora. De exemplu, la categoria „cuvinte cu sens contrar”, un indiciu ar putea fi ceva de genul:

Un element de fixare pentru a prinde ceva sau pentru a îndoi, a deforma şi a elibera brusc ceva, prin intermediul căldurii sau al presiunii.

Concurentul-algoritm ar trebui să înveţe să treacă prin câteva niveluri pentru a obţine răspunsul corect, respectiv: „Ce este o «chingă»?” Mai întâi, Watson ar trebui să înţeleagă suficient de bine limbajul pentru a desprinde înţelesul din întrebare şi a-şi da seama că „element de fixare”, „a prinde”, „a îndoi”, „a deforma” şi

„a elibera brusc” sunt elemente separate care compun indiciul. Aceasta, în sine, este o provocare uriaşă pentru un algoritm.

Şi acela nu reprezenta decât primul pas! După aceea, Watson trebuia să caute potenţiali candidaţi care să se potrivească fiecărui indiciu în parte. „Element de fixare” ar putea evoca tot felul de posibile răspunsuri: „clemă”, „buton”, „agrafă” şi „legătură”, dar şi „chingă”, de exemplu. Watson trebuia să analizeze pe rând fiecare posibilitate şi să calculeze cum se potrivea aceasta cu restul indiciilor. Aşadar, cu toate că este puţin probabil să găseşti dovezi că „agrafa” s-ar putea asocia cu indicii precum „a îndoi” şi „a deforma”, cuvântul „chingă” sigur se asociază cu ele, ceea ce creşte siguranţa lui Watson că acesta este un posibil răspuns. În final, odată ce toate elementele au fost combinate, Watson trebuia să-şi susţină părerea cu fapte, alegând un singur răspuns.

Provocarea de a juca *Jeopardy* este mai banală decât aceea de a diagnostica o afecţiune, însă ambele necesită câteva mecanisme logice comune. Imaginaţi-vă că vă duceţi la doctor, acuzând o scădere în greutate neintenţionată, dureri de stomac şi (ca să fie tacâmul complet!) ceva arsuri în capul pieptului. Ca şi atunci când joci *Jeopardy*, provocarea o constituie să găseşti posibile diagnostice (răspunsuri) care ar putea explica simptomele (indiciile), să cauţi dovezi suplimentare pentru fiecare în parte şi să-ţi reiterezi încrederea într-un anumit răspuns pe măsură ce devin disponibile mai multe informaţii. Doctorii numesc asta „diagnostic diferenţial”. Matematicienii o numesc „inferenţă bayesiană”[43]

Chiar şi după crearea încununată de succes a lui Watson, câştigătorul concursului de cultură generală, reconfigurarea sa ca geniu medical nu a fost o sarcină uşoară. Când, însă, IBM şi-a declarat public intenţia de a pătrunde în domeniul sănătăţii, nu s-a sfiit să facă

43 Mai multe despre Bayes veţi afla în capitolul „Automobilele autonome” (n. a.).

promisiuni măreţe. A anunţat întreaga lume că misiunea supremă a lui Watson era „eradicarea cancerului” 35 şi l-a angajat pe celebrul actor John Hamm să îl laude, numindu-l „una dintre cele mai puternice unelte create de specia umană”.

Aceasta este, cu siguranţă, viziunea unei utopii medicale menite să ne inspire pe toţi. Cu excepţia faptului că – aşa cum probabil deja ştiţi – Watson nu s-a prea ridicat la nivelul aşteptărilor.

Mai întâi, un contract prestigios încheiat cu Centrul pentru studierea cancerului „M.D. Anderson” al Universităţii din Texas a fost reziliat în 2016. Se zvoneşte că nici după ce au plătit 62 de milioane de dolari pe tehnologie36 şi au petrecut patru ani lucrând la el, Watson nu a reuşit să facă mai mult decât teste de verificare strict controlate. Apoi, la sfârşitul lui septembrie 2017, o investigaţie a site-ului de ştiri din domeniul sănătăţii STAT a descoperit că Watson „încă se chinuia cu etapa de bază a învăţării despre diferitele forme de cancer”.37

Îngrozitor, ce mai!

Cinstit vorbind, nu au fost numai veşti proaste. În Japonia, Watson reuşise să diagnosticheze o femeie care avea o formă rară de leucemie, atunci când medicii nu izbutiseră acest lucru.38 Iar analizele făcute de Watson au dus la descoperirea a cinci gene legate de boala neuronilor motorii, denumită şi scleroza laterală amiotrofică.39

În linii mari, însă, programatorii de la IBM nu au prea reuşit să respecte promisiunile departamentului lor de marketing, care fusese atât de entuziast.

Este greu să nu simţi o oarecare simpatie faţă de oricine încearcă să construiască o astfel de maşină. Teoretic, este posibil să concepi una care poate să diagnosticheze diverse afecţiuni (şi chiar să le ofere pacienţilor planuri de tratament de bun-simţ), acesta fiind un scop admirabil în sine. Numai că este şi foarte greu. Mult mai greu decât să concurezi la *Jeopardy* şi mult, mult mai greu decât să recunoşti celulele canceroase într-o

imagine.

O maşină de diagnosticare multifuncţională ar putea să pară doar un simplu pas logic de la algoritmii care se bazau pe imagini pentru depistarea cancerului pe care i-am întâlnit mai devreme, însă aceia aveau un mare avantaj: examinau celule reale care ar fi putut să cauzeze o anumită problemă. O maşină de diagnosticare, în schimb, primeşte informaţii în condiţiile în care din fondul problemei sunt eliminate câteva etape. Poate că un pacient simte furnicături provocate de un spasm muscular produs de prinderea unui nerv din cauza ridicării unei greutăţi prea mari. Sau poate că o persoană are sânge în scaun din cauza hemoroizilor apăruţi în urma unei constipaţii provocate de alimentaţia necorespunzătoare. Un algoritm (sau un medic) trebuie să ia un singur simptom şi să refacă traseul până când ajunge la diagnosticul corect. Asta trebuia să facă şi Watson. Este o sarcină incredibil de dificilă.

Şi mai există şi alte probleme, pe deasupra!

Vă mai aduceţi aminte de reţeaua neurală care deosebea un câine de un lup? Antrenarea sa fusese uşoară. Tot ce avuseseră de făcut programatorii fusese să găsească un teanc de fotografii etichetate „lup" sau „câine" şi să le introducă în calculator. Setul de date era simplu şi lipsit de ambiguităţi. „Numai că – aşa cum a declarat pentru *MIT Technology Review* Thomas Fuchs, specialist în patologie informatică[44] într-un domeniu specializat al medicinei ai putea avea nevoie de specialişti instruiţi timp de decenii care să nu facă nimic altceva decât să eticheteze corect informaţiile introduse în computer". 40

Această problemă ar putea fi depăşită cu ajutorul unei întrebări foarte specializate (de exemplu, sortarea lamelelor cu ţesut mamar care ajung la patologie în „complet benigne" şi „foarte maligne"). Însă o maşină de diagnosticare universală, cum este Watson, ar avea nevoie să înţeleagă practic orice afecţiune posibilă. Ar fi nevoie de

44 Computational pathologist (engl.) (n. tr.).

o armată de operatori umani incredibil de bine calificaţi, pregătiţi să introducă în calculator informaţii despre diferiţi pacienţi şi caracteristicile lor specifice pe o perioadă foarte lungă de timp. Şi, dacă tot vorbim la modul general, patologii aceştia tind să aibă şi alte lucruri de făcut – să salveze vieţi, de exemplu.

Ajungem, astfel, la cea din urmă problemă. Cea care este cel mai dificil de depăşit.

PROBLEMA DATELOR

Tamara Mills era doar un bebeluş când părinţii ei au observat pentru prima oară că era ceva în neregulă cu felul în care respira. Când împlinise nouă luni, doctorii o diagnosticaseră ca fiind bolnavă de astm – o boală care afectează 5, 4 milioane de oameni în Marea Britanie şi 25 de milioane în SUA.41 Cu toate că Tamara era mai tânără decât cei mai mulţi suferinzi de astm, simptomele sale au fost perfect controlabile în primii ani de viaţă şi fetiţa a crescut cam ca orice alt puşti cu această boală, petrecându-şi copilăria jucându-se pe malul mării, în nordul Angliei (dar având tot timpul un inhalator la îndemână).

La opt ani, Tamara s-a îmbolnăvit de o formă urâtă de gripă porcină. De atunci, o infecţie respiratorie s-a succedat după alta. Câteodată, în timpul atacurilor de astm, i se învineţeau buzele. Dar indiferent de câte ori Tamara şi mama ei se duceau la doctorii de la spitalul din apropiere, indiferent cât de des se plângeau părinţii ei că îşi termina rezerva de inhalatoare mai repede decât i se puteau prescrie42, niciunul dintre doctori nu a trimis-o la un specialist.

Familia şi profesorii Tamarei, pe de altă parte, şi-au dat seama că lucrurile se agravau. După două atacuri aproape fatale care au dus la internarea fetei în spital, aceasta a fost scutită de orele de educaţie fizică de la şcoală. Când scările interioare au început să fie prea greu de urcat pentru ea, Tamara s-a dus să locuiască la bunici, într-un bungalow.

Pe 10 aprilie 2014, Tamara căzut victimă unei noi infecţii respiratorii. În noaptea aceea, bunicul ei a găsit-o chinuindu-se să respire. A chemat o ambulanţă şi a făcut tot ce i-a stat în putinţă să o ajute, cu două inhalatoare şi un tub de oxigen. Starea ei a continuat să se deterioreze. Tamara a murit câteva ore mai târziu, la numai 13 ani.

Astmul nu este, de obicei, o afecţiune fatală, dar, cu toate acestea, 1 200 de persoane de pe teritoriul Marii Britanii, dintre care 26 sunt copii, mor din cauza lui în fiecare an.43 Se estimează că două treimi dintre aceste decese ar putea fi prevenite – aşa cum a fost şi cel al Tamarei. Numai că prevenirea depinde în întregime de identificarea semnelor de avertizare, în baza cărora se şi intervine.

Pe durata celor patru ani care au culminat cu cel din urmă atac, care i-a fost fatal, Tamara a fost la consultaţii la medicii de la spitalul din localitate de nu mai puţin de 47 de ori. Era evident că tratamentul ei nu dădea rezultate, dar, de fiecare dată când se ducea la un specialist în sănătate, acesta nu trata decât problema cea mai urgentă. Niciunul nu a luat în considerare imaginea de ansamblu. Niciunul nu a remarcat tiparul care reieşea la vizitele medicale, niciunul nu a observat că starea fetei se deteriora constant, niciunul nu a sugerat că era momentul să încerce ceva nou.44

Pentru toate acestea exista un motiv. Indiferent dacă vă vine sau nu să credeţi (şi oricine locuieşte în Anglia probabil că va crede), pentru NHS[45], serviciul de sănătate naţional din Marea Britanie, corelarea evidenţelor serviciilor medicale prestate nu este o procedură standard. Dacă te afli într-un spital din sistemul naţional de sănătate, medicii de acolo nu vor şti nimic despre vizitele pe care le-ai făcut la medicul de familie. Multe evidenţe sunt încă ţinute pe hârtie, ceea ce înseamnă că modul în care trec de

45 National Health Service (engl.), Serviciul Medical Naţional, sistemul public de asistenţă medicală din Regatul Unit al Marii Britanii şi Irlandei de Nord, care nu are o reputaţie bună în ceea ce priveşte protejarea datelor personale ale pacienţilor şi angajaţilor săi (n. red.).

la un medic la altul nu s-a schimbat de zeci de ani. Este unul dintre motivele pentru care sistemul de sănătate britanic deţine titlul nu prea lăudabil de cel mai mare cumpărător de faxuri din lume.45

Oricât de nebunesc ar suna asta, noi, britanicii, nu suntem singurii. SUA au o puzderie de medici particulari şi mari reţele de spitale care nu sunt conectate între ele; pe când alte ţări, precum Germania, au început să introducă dosarele electronice de sănătate ale pacienţilor, deşi acestea nu reprezintă, nici pe departe, ceva obişnuit peste tot în lume. Pentru Tamara Mills, lipsa unui istoric medical unic conectat la întregul sistem a însemnat că era imposibil pentru orice medic individual să înţeleagă pe deplin severitatea stării ei. Orice soluţionare a acestei deficienţe va veni, din păcate, prea târziu pentru Tamara, dar rămâne o provocare uriaşă pentru viitorul sistemului de sănătate. O maşină cum este Watson ar putea ajuta la salvarea unei mulţimi de Tamare, însă va reuşi să găsească anumite tipare în informaţiile pe care le deţine numai dacă acestea sunt colectate, centralizate şi conectate.

Există un contrast izbitor între seturile de date complete şi detaliate deţinute de brokerii de informaţii şi seturile de date restrânse şi necorelate din sistemul de sănătate. Pentru moment, informaţiile medicale sunt vraişte. Chiar şi atunci când istoricul nostru medical detaliat este stocat într-un singur loc (ceea ce adesea nu este cazul), informaţiile însele pot lua atâtea forme, încât este practic imposibil ca acestea să fie corelate într-un mod utilizabil de către un algoritm. Sunt de luat în considerare radiografii, trebuie incluse rapoarte şi fişe, reţete, observaţii; iar lista poate continua. Apoi mai este problema modului în care sunt înregistrate informaţiile scrise. Trebuie să înţelegi acronimele şi abrevierile, să descifrezi scrisul de mână, să identifici posibilitatea apariţiei unor erori umane. Toate acestea, înainte de a ajunge măcar la simptome. Oare această persoană s-a referit la faptul că are temperatură când a spus că suferă

de o „răceală”? Sau „răceală” înseamnă că tuşeşte? Pe această persoană chiar „o omoară” durerile de stomac? Sau simte doar puţină durere? Ideea este că medicina este foarte, foarte complicată şi fiecare nivel de complexitate face ca informaţiile să fie şi mai greu de priceput pentru un computer.46

Cei de la IBM nu sunt singurii giganţi economici *blue chip*[46] care au avut de furcă din cauza problemei reprezentate de bazele de date dezordonate şi nestructurate ale sistemului de sănătate. În 2016, Deep Mind, divizia dedicată dezvoltării IA a companiei Google, a semnat un contract cu Royal Free NHS Trust din Londra. Deep Mind a primit acces la informaţiile medicale deţinute de trei spitale municipale, în schimbul unei aplicaţii care i-ar fi ajutat pe medici să identifice insuficienţele renale acute. Iniţial, se intenţiona utilizarea algoritmilor inteligenţi de învăţare care să vină în sprijinul serviciilor medicale, însă cercetătorii au descoperit că trebuiau să-şi limiteze ambiţiile şi să opteze pentru ceva mult mai simplu, deoarece informaţiile nu erau suficient de bune pentru îndeplinirea ţelurilor iniţiale.

Dincolo de provocările de natură strict practică, această colaborare dintre Deep Mind şi NHS a dat naştere unei probleme încă şi mai controversate. Cercetătorii nu promiseseră decât să îi avertizeze pe medici despre prezenţa afecţiunilor renale, numai că Royal Free nu dispunea de un set de date despre rinichi pe care să li-l ofere. În schimb, Deep Mind a primit acces deplin la *toate* informaţiile înregistrate în sistem, adică la antecedentele medicale ale unui număr de 1, 6 milioane de pacienţi pe o perioadă de cinci ani.

Teoretic, o cantitate aşa de mare de informaţii ar putea ajuta la salvarea a nenumărate vieţi. Insuficienţele

46 Expresia blue chip, derivată din jargonul jucătorilor de pocher (care folosesc jetoane de diferite culori, dintre care cele albastre sunt cele mai valoroase), este o etichetă folosită pentru companiile publice considerate cele mai solide, mai eficiente şi mai stabile de pe piaţă (n. tr.).

renale acute ucid o mie de persoane lunar, iar deţinerea unor informaţii care se întindeau pe o perioadă atât de îndelungată putea să ajute Deep Mind să identifice tendinţe importante referitoare la evoluţia acestor afecţiuni. Mai mult decât atât, din moment ce bolile de rinichi se întâlnesc mai frecvent la persoane care suferă şi de alte maladii, un set de date cuprinzător ar uşura mult depistarea simptomelor şi a diverselor relaţii dintre variatele afecţiuni din prezent ale pacienţilor şi starea lor de sănătate viitoare.

În loc să fie întâmpinată cu entuziasm, anunţarea proiectului a stârnit însă un adevărat scandal. Şi nu fără temei. A-i oferi lui Deep Mind accesul la *toate* informaţiile înregistrate de spitale însemna exact asta: *că avea acces la tot*. Compania Alphabet Inc. (care deţine şi Google) a aflat cine a fost internat în spital şi la ce dată. Cine a venit în vizită la pacienţi în perioada internării acestora. A aflat rezultatele rapoartelor patologice şi ale radiografiilor; cine a făcut întreruperi de sarcină, cine a suferit de depresie, până şi cine a fost diagnosticat cu HIV. Şi ştiţi ce este cel mai grav? Pacienţilor înşişi nu li s-a cerut niciodată consimţământul, nu li s-a oferit posibilitatea de a fi excluşi din studiu şi nici măcar nu li s-a spus că urmau să facă parte din acesta.47

Merită amintit că i s-a interzis companiei Google să utilizeze aceste informaţii în orice alt context al activităţii sale comerciale. Şi, cinstit vorbind, Google deţine în privinţa securităţii informaţiilor un palmares superior celui al NHS, ale cărui spitale au fost aduse în impas în 2017 de un virus informatic de tip *ransomware* lansat de Coreea de Nord, pentru că NHS încă mai folosea Windows XP.48 Chiar şi aşa, este un motiv serios de îngrijorare faptul că o companie deja extrem de puternică, lider mondial în domeniul tehnologiei, a ajuns să aibă acces şi la informaţiile medicale ale unei persoane.

PROBLEME DE CONFIDENŢIALITATE

Haideţi să fim sinceri: Google nu duce lipsă de

informaţii personale, intime chiar, despre fiecare dintre noi. Dar parcă simţim instinctiv ceva diferit – extrem de confidenţial – atunci când vine vorba despre informaţiile noastre medicale.

Pentru oricine se bucură de o stare de sănătate bună s-ar putea să nu fie atât de evident de ce: în definitiv, dacă ai avea de ales între a-ţi face publice informaţiile medicale sau istoricul navigării pe internet, ce ai prefera? Eu ştiu că l-aş alege, fără să stau pe gânduri, pe primul şi probabil că mulţi ar proceda la fel. Nu pentru că aş avea ceva deosebit de interesant de ascuns. Dar una este doar un instantaneu banal al fiinţei mele biologice, pe când cealaltă este o fereastră deschisă spre personalitatea mea.

Chiar dacă informaţiile despre starea de sănătate ar putea fi mai puţin stânjenitoare, Timandra Harkness, autoarea cărţii *Big Data*: *Does Size Matter*? şi prezentatoare a programului *Future Proofing* de la BBC Radio 4, susţine că acestea continuă să reprezinte un caz aparte:

Înainte de orice, informaţiile despre starea de sănătate a multor persoane conţin o naraţiune a vieţii lor. Una din trei femei din Marea Britanie, de exemplu, a făcut o întrerupere de sarcină şi ar putea există persoane din viaţa lor care să nu fie la curent cu acest lucru.

Timandra subliniază şi faptul că istoricul medical nu este important numai pentru tine:

Dacă o persoană deţine informaţiile tale genetice, atunci ştie câte ceva şi despre părinţii, fraţii şi copiii tăi. Iar când aceste informaţii sunt făcute publice, nu ai cum să mai scapi de asta. Nu-ţi poţi schimba structura biologică şi nici nu o poţi nega. Dacă examinează cineva un eşantion de ADN care îţi aparţine, nu poţi schimba acest lucru. Poţi face operaţii estetice la faţă, poţi purta mănuşi pentru a-ţi ascunde amprentele, însă ADN-ul rămâne întotdeauna acolo. El este asociat cu persoana ta pentru totdeauna.

De ce este important acest lucru? Timandra mi-a

relatat despre un *focus grup*[47] pe care l-a condus în anul 2013, în care oameni obişnuiţi erau întrebaţi ce îi preocupa cel mai mult în privinţa informaţiilor lor medicale:

În ansamblu, oamenii nu erau îngrijoraţi de faptul că informaţiile lor ar putea fi piratate sau furate. Îi preocupa faptul că s-ar putea face supoziţii despre ei ca grup, care să fie apoi proiectate asupra lor ca indivizi.

Mai presus de orice, erau preocupaţi de modul în care aceste informaţii ar fi putea fi folosite împotriva lor:

Să presupunem că cineva şi-a asociat cârdul de fidelitate emis de un supermarket la istoricul medical. Această persoană s-ar putea duce să facă o operaţie de şold, iar medicul i-ar spune, de exemplu: „O, îmi pare rău, am constatat că aţi cumpărat o mare cantitate de pizza sau că aţi cumpărat o mulţime de ţigări, motiv pentru care mă tem că vom fi nevoiţi să vă trimitem pe ultimul loc al listei de aşteptare”.

Este o temere foarte îndreptăţită în Marea Britanie, unde anumite spitale publice ale NHS, secătuite cum sunt de fonduri, acordă deja prioritate nefumătorilor când vine vorba de realizarea unor operaţii de şold şi genunchi.49 Şi există multe ţări din lume în care pacienţilor obezi le pot fi respinse cererile de asigurare sau de tratament medical.50

Avem, aici, un fel de dilemă morală. Oamenii, ca specie, ar avea enorm de câştigat dacă le-ar permite algoritmilor să aibă acces direct la dosarele medicale algoritmi. Watson nu trebuie neapărat să rămână o simplă fantezie. Însă, pentru a-l transforma în realitate, va trebui să ne oferim informaţiile medicale unor companii îndeajuns de bogate, încât să ne poarte prin corvoada încercărilor care se interpun între noi şi acel doctor electronic fermecat. Iar, ori de câte ori renunţăm de bunăvoie la confidenţialitate, ne jucăm periculos cu focul,

47 Metodă de cercetare bazată pe discuţii colective libere, care explorează o anumită întrebare sau un ansamblu de întrebări în cadrul unui grup bine definit (demografic, sociologic etc.) (n. tr.).

fiind în pericol ca datele personale să ne fie eventual compromise, furate sau folosite împotriva noastră. Dumneavoastră sunteţi pregătit să vă asumaţi acest risc? Aveţi suficientă încredere în algoritmii aceştia şi în beneficiile lor ca să vă sacrificaţi intimitatea?

Sau, dacă se va ajunge aici, o să vă pese foarte tare de ea?

CÂND MOŞTENIREA GENETICĂ TE DĂ DE GOL

Francis Galton a fost un statistician şi genetician din epoca victoriană, unul dintre cei mai remarcabili oameni ai generaţiei sale şi văr pe jumătate cu Charles Darwin. Multe dintre ideile sale au avut un efect profund asupra ştiinţei moderne şi, nu în ultimul rând, opera sa a pus bazele statisticii moderne. Pentru aceasta îi datorăm lui Galton toată recunoştinţa noastră. (Din păcate, el a fost destul de activ şi în mişcarea eugenică, aflată pe atunci în plin avânt, iar pentru asta, în mod cert, nu trebuie să îi fim recunoscători.)

Galton a vrut să studieze caracteristicile umane prin intermediul informaţiilor; chiar şi atunci, îşi dădea seama că era nevoie de o mulţime de date ca să afli ceva cu adevărat demn de interes. Dar mai ştia că oamenii nutreau o curiozitate insaţiabilă cu privire la propriul corp. Şi-a dat seama că – atunci când era stimulat în maniera potrivită – apetitul de nepotolit al oamenilor faţă de evaluarea lor de către un expert putea să depăşească nevoia de confidenţialitate. Mai mult decât atât, adesea erau dispuşi să plătească pentru plăcerea de a-şi satisface acest apetit.

Aşa se face că în 1884, când s-a organizat la Londra, sub patronajul reginei Victoria, o imensă expoziţie pentru a sărbători progresele Marii Britanii în domeniul sănătăţii, Galton şi-a văzut visul cu ochii. Şi-a instalat un stand la expoziţie pe cheltuiala sa – stand pe care l-a numit „Laborator antropometric” – în speranţa că va găsi câteva persoane printre milioanele de vizitatori dornice să plătească nişte bani pentru a fi măsurate.

A găsit chiar mai multe. Clienţii au format o coadă şi

s-au îmbulzit la uşă, dornici să plătească un fleac de trei penny de persoană pentru a intra în laborator. Odată intraţi, puteau să-şi compare aptitudinile prin intermediul unor instrumente special concepute în acest scop, verificându-şi, printre altele, acuitatea vederii, forţa cu care puteau să tragă sau să strângă mâna şi viteza cu care trăgeau un pumn.51 Laboratorul a fost atât de apreciat, încât Galton a trebuit să evalueze câte două persoane odată. (Şi-a dat repede seama însă că era mai bine să separe părinţii de copii în timpul testării pentru a evita pierderea de timp din cauza orgoliilor rănite. Într-un articol de ziar pe care l-a scris după eveniment, Galton comenta: „Celor vârstnici nu le plăcea să fie depăşiţi de cei tineri şi insistau să repete testele.”) 52

Indiferent cât de bine sau cât de rău se descurca fiecare persoană, rezultatele pe care le obţinuse erau mâzgălite pe un cartonaş alb pe care îl putea păstra ca amintire. Adevăratul câştigător a fost, însă, Galton. Acesta a părăsit expoziţia cu o copie completă a tuturor informaţiilor – un set valoros de măsurători biometrice ale unui număr de 9 337 de persoane – şi cu un profit frumuşel pe deasupra.

Dacă trecem repede peste următorii 130 de ani, veţi putea observa unele asemănări cu moda actuală a testărilor genetice. La un preţ „de chilipir”, de numai 149 de lire sterline, puteţi trimite o probă de salivă companiei de genomică şi biotehnologie 23 and Me, pentru a primi în schimb un raport al trăsăturilor genetice, incluzând răspunsuri la întrebări de genul: ce tip de ceară de urechi produci? *Tu* ai cumva genele care duc la apariţia unei monosprâncene? 53 Sau genele care te fac să strănuţi când te uiţi la soare? 54 Precum şi la unele întrebări mai serioase: sunteţi predispusă la cancer mamar? Aveţi o predispoziţie genetică pentru maladia Alzheimer? 55

Între timp, compania a strâns, în acest mod ingenios, o uriaşă bază de date cu informaţii genetice, care numără acum milioane de probe. Este exact ce au făcut şi giganţii

internetului, numai că lor, în loc să le dăm ADN-ul nostru ca parte a târgului, le oferim informaţiile cele mai personale pe care le deţinem. Rezultă de aici o bază de date de pe urma căreia avem cu toţii de câştigat. Aceasta constituie un bun de o valoare inestimabilă, căci ne va ajuta să înţelegem mai bine genomul uman. Oameni de ştiinţă, companii farmaceutice şi organizaţii non-profit din întreaga lume aşteaptă la rând să încheie parteneriate cu 23 and Me pentru a depista anumite tipare în informaţiile pe care aceasta le deţine – atât eu, dar şi fără ajutorul algoritmilor – în speranţa găsirii unui răspuns la marile întrebări care ne afectează pe toţi: care sunt cauzele ereditare ale diverselor afecţiuni? S-ar putea inventa medicamente noi, care să trateze persoanele care suferă de o anumită boală? Există o cale mai bună de tratare a maladiei Parkinson?

Setul de date este valoros şi într-un sens mai concret. Cu toate că fiecare cercetare întreprinsă aduce un beneficiu imens pentru societate, 23 and Me nu face asta din bunătate şi altruism. Dacă vă daţi consimţământul (şi 80% dintre clienţi o fac), compania va vinde mai departe, pentru un profit frumuşel, o versiune anonimizată a informaţiilor dumneavoastră genetice partenerilor de cercetare mai susmenţionaţi.56 Banii astfel câştigaţi nu sunt vreun bonus binevenit pentru companie, ci fac parte efectiv din planul ei de afaceri. Un membru din consiliul de administraţie al 23 and Me a declarat pentru *Fast Company*: „Jocul pe termen lung în acest caz nu este să facem bani din vânzarea de kituri, deşi kiturile sunt indispensabile pentru obţinerea informaţiilor de bază”. Iată ceva demn de reţinut atunci când comandaţi un raport genetic comercial: nu dumneavoastră folosiţi produsul. Dumneavoastră *sunteţi* produsul.57

V-aş face un avertisment: eu, una, nu m-aş încrede prea mult în promisiunile lor de păstrare a anonimităţii. În 2005, un tânăr58 conceput cu ajutorul unui donator de spermă rămas anonim a reuşit să-l găsească şi să-l

identifice pe tatăl său biologic prin trimiterea unei probe de salivă la analiză şi prin urmărirea indiciilor din propriul cod ADN.59 Iar în 2013, un grup de oameni de ştiinţă a demonstrat, într-o lucrare foarte cunoscută, că milioane de oameni ar putea fi identificaţi după propriile gene, folosindu-se doar un calculator personal şi câteva căutări inteligente pe internet.60

Şi mai există un motiv pentru care s-ar putea să nu vă doriţi ca ADN-ul dumneavoastră să fie înscris în vreo bază de date. Cu toate că există legi în vigoare care să protejeze oamenii împotriva celor mai urâte feluri de discriminare genetică – aşa încât să nu ne îndreptăm chiar spre un viitor în care un Beethoven sau un Stephen Hawking să fie judecaţi mai degrabă după predispoziţiile lor genetice decât după talent aceste legi nu se aplică şi în cazul asigurărilor de viaţă. Nimeni nu te poate sili să faci un test ADN dacă nu doreşti acest lucru, însă în SUA asigurătorii au dreptul să te întrebe dacă ai făcut deja vreun test care îţi calculează riscul de a dezvolta o anumită afecţiune, cum ar fi maladiile Parkinson, Alzheimer sau cancerul mamar, şi pot să refuze încheierea unei asigurări de viaţă dacă răspunsul nu-i satisface. Iar în Marea Britanie asigurătorilor li se permite să ţină cont de rezultatele testelor genetice pentru depistarea bolii Huntington (dacă suma acoperită depăşeşte 500.000 de lire sterline).61 Bineînţeles că aţi putea încerca să minţiţi şi să susţineţi că nu aţi făcut niciodată un astfel de test, dar asta ar duce la invalidarea poliţei de asigurare. Singura cale de a evita acest gen de discriminare este să nu faceţi niciodată testul. Uneori, ignoranţa chiar poate fi o binecuvântare.

Adevărul este că niciun set de date nu ar fi mai valoros pentru înţelegerea stării noastre de sănătate decât genomurile secvenţiate a milioane de persoane. Cu toate acestea, nu cred că voi face un test genetic în viitorul apropiat. Însă (din fericire pentru societatea noastră, în ansamblu) milioane de oameni îşi *oferă* informaţiile de

bunăvoie. La ultima numărătoare, 23 and Me avea peste două milioane de clienţi genotipizaţi62, iar My Heritage, Ancestry.com – chiar şi proiectul *Genografic*[48] al National Geographic – aveau şi mai multe milioane. Aşadar, poate că aici nu există niciun fel de dilemă morală. În definitiv, piaţa şi-a spus cuvântul: s-ar putea să nu merite ca în schimbul confidenţialităţii informaţiilor tale să ţi se ofere şansa de a contribui la binele public, dar să descoperi că eşti 25% viking merită cu prisosinţă.[49]

Bine, bine, recunosc: am glumit! Nimeni nu se aşteaptă să se gândească, totuşi, cineva în primul rând la marile provocări ale viitorului sistemului nostru de sănătate atunci când încearcă să ia decizia dacă să expedieze sau nu o probă de salivă pentru un test genetic. Este de înţeles dacă oamenii nu fac acest lucru – avem priorităţi diferite pentru noi înşine ca indivizi şi pentru umanitate în ansamblu.

Dar asta ne trimite la un ultim aspect important. Dacă o maşină de diagnosticare capabilă să recomande tratamente medicale *poate* fi totuşi construită, cui ar trebui ea să-i servească? Individului sau umanităţii? Pentru că vor fi momente în care ar putea fi nevoită să aleagă.

Imaginaţi-vă, de exemplu, că vă duceţi la doctor din cauză că vă supără teribil tuşea. Probabil că vă veţi însănătoşi de la sine, dar, dacă v-ar consulta o maşină – ca pacient –, ea ar putea dori să vă trimită să vă faceţi o radiografie şi un test de sânge, doar aşa, pentru siguranţă. Şi probabil că v-ar prescrie antibiotice, dacă i le-aţi solicita. Chiar dacă asta v-ar scurta suferinţa cu doar câteva zile, dacă unicul obiectiv ar fi propria sănătate şi starea de bine, algoritmul ar putea considera că a meritat să vă prescrie tot acest tratament.

Dacă, însă, s-ar construi o maşină care să servească

48 Un proiect de cercetare a evoluţiei şi migraţiei speciei umane lansat la 13 aprilie 2005 de Societatea National Geographic (n. tr.).

49 De fapt, nimeni nu îşi poate da seama dacă cineva este sau nu viking, aşa cum m-a informat pe larg bunul meu prieten, geneticianul Adam Rutherford. Am introdus această informaţie mai mult ca să-l

întregii populaţii, aceasta ar fi mult mai preocupată de problemele legate de rezistenţa la antibiotice. Atât timp cât nu v-aţi afla într-un pericol iminent, disconfortul dumneavoastră de moment ar păli în importanţă, iar algoritmul v-ar prescrie antibiotice numai dacă acest lucru ar fi absolut necesar. Un astfel de algoritm ar fi preocupat şi de irosirea resurselor sau ar fi conştient de lungile liste de aşteptare, astfel că nu v-ar trimite să faceţi investigaţii suplimentare decât dacă aţi avea şi alte simptome care să denote ceva mai grav. Sincer vorbind, probabil că v-ar spune să luaţi o aspirină şi să nu vă mai plângeţi atât de milă.

În mod similar, o maşină care ar lucra în beneficiul tuturor ar putea să facă din „salvarea a cât mai multor vieţi cu putinţă” o prioritate atunci când ar trebui să decidă cine să primească un transplant de organe. Ceea ce ar putea produce un plan de tratament diferit de cel al unei maşini care ar ţine cont doar de interesul dumneavoastră personal.

O maşină care ar lucra pentru NHS sau pentru o companie de asigurări ar putea încerca să micşoreze costurile, dac-ar fi posibil, pe când una concepută să servească interesele unei companii farmaceutice ar putea avea ca obiectiv să promoveze utilizarea unui anumit medicament în locul altuia.

Situaţia medicinei este, cu siguranţă, mai lipsită de tensiuni decât exemplele din sistemul de justiţie penală din capitolul anterior. Aici nu avem apărare şi acuzare. Toţi cei din sistemul de sănătate au acelaşi deziderat: să facă pacientul să se simtă mai bine. Dar chiar şi aici, fiecare parte implicată în proces are un set de obiective puţin diferit.

În oricare aspect al vieţii s-ar introduce un algoritm, va trebui să se creeze întotdeauna un fel de echilibru. Între confidenţialitate şi binele public. Între individ şi populaţie, între diverse dificultăţi şi priorităţi. Nu este uşor să găseşti o cale prin încâlceala de motivaţii, chiar şi atunci când la

capătul drumului aşteaptă recompensa sigură a unui sistem de sănătate mai bun pentru toată lumea.

Dar este cu atât mai greu atunci când motivaţii concurente sunt ascunse vederii; când beneficiile unui algoritm sunt exagerate, iar riscurile – disimulate; sau când trebuie să te întrebi ce ţi se cere să crezi şi cine obţine un profit de pe urma încrederii tale într-un anumit lucru.

Capitolul 5
AUTOMOBILELE AUTONOME

Soarele abia se ivise la orizont în dimineaţa zilei de 13 martie 2004, dar barul Slash X din inima deşertului Mojave deja gemea de oameni.1 Localul se afla la periferia Barstowului, un orăşel situat între Los Angeles şi Las Vegas, situat în apropierea locului unde a fost filmată Urna Thurman ieşind dintr-un sicriu în filmul *Kill Bill II.2* Este un bar îndrăgit de cowboy şi de cei cărora le plac vehiculele de teren, dar, în acea dimineaţă de primăvară, reuşise să atragă un cu totul alt gen de clienţi. Stadionul improvizat care fusese construit afară, în praf, era înţesat de ingineri nebuni, de spectatori entuziasmaţi şi de pasionaţi nesăbuiţi, care împărtăşeau acelaşi vis: să fie primii oameni de pe planetă care asistau la câştigarea unei curse de către un vehicul fără şofer.

Cursa fusese organizată de US Defence Advanced Research Projects Agency*, DARPA (care mai era poreclită şi „departamentul de ştiinţă nebună al Pentagonului”).3 Agenţia era interesată de ceva timp de vehiculele fără şofer, şi pe bună dreptate: bombele artizanale şi atacurile ţintite asupra vehiculelor militare constituiau o cauză majoră a deceselor pe câmpul de luptă. Cu ceva timp în urmă, în acel an, îşi anunţase intenţia de a face autonome o treime din vehiculele forţelor militare la sol ale SUA până în anul 2015.4

— Agenţia pentru Proiecte de Cercetare Avansată în domeniul Apărării a SUA (n.tr.).

Până în acel moment, progresul fusese lent şi

costisitor. DARPA cheltuise în jur de o jumătate de miliard de dolari în peste 20 de ani, finanţând munca de cercetare din universităţi şi companii, în speranţa realizării acestei ambiţii.5 Şi, la un moment dat, mai-marilor DARPA le-a venit o idee ingenioasă: ce-ar fi să organizeze un concurs? Ar invita în mod deschis toate persoanele interesate din ţară, care să-şi conceapă propriile vehicule fără şofer şi să le pună să concureze una împotriva celeilalte pe o pistă de lungă distanţă, oferind învingătorului un premiu de un milion de dolari.6 Era primul eveniment de acest fel din lume şi un mod rapid şi ieftin de a oferi DARPA un avantaj în îndeplinirea obiectivului său.

Traseul cursei a fost amenajat pe ceva mai mult de 228 de kilometri, iar DARPA nu l-a conceput deloc pentru a fi uşor. Participanţii trebuiau să se confrunte cu pante abrupte, bolovani, denivelări, şanţuri, porţiuni de teren accidentat şi cactuşi ciudaţi. Automobilele fără şofer trebuiau să parcurgă porţiuni cu zgură care, pe alocuri, aveau doar vreun metru şi ceva lăţime. Cu două ore înainte de start, organizatorii au oferit fiecărei echipe un CD cu coordonate GPS.7 Acestea reprezentau 2000 de puncte de trecere răspândite de-a lungul traseului asemenea firmiturilor de pâine – doar atât cât să le ofere vehiculelor o idee generală asupra direcţiei de mers, dar nu suficiente pentru a le ajuta să treacă peste obstacolele care le aşteptau.

Provocarea era dificilă, însă 106 echipe curajoase s-au înscris la concurs în acel prim an, iar 15 au trecut de etapele de calificare, iar vehiculele lor au fost considerate suficient de sigure, încât să li se îngăduie să ajungă pe pistă. Printre acestea se numărau automobile care semănau cu buggyurile destinate mersului pe nisip, vehicule care semănau cu nişte camioane gigantice şi altele care păreau, mai degrabă, tancuri. Circulau zvonuri precum că un concurent îşi ipotecase casa ca să-şi construiască automobilul fără şofer, pe când un altul lipise două plăci de surf pe capotă ca să o facă să iasă în

evidenţă. La întrecere a participat chiar şi o motocicletă cu autobalans.8

În dimineaţa cursei, un şir de maşinării improvizate şi care păreau cam şubrede s-au adunat la Slash X, alături de câteva mii de spectatori. Fără şoferi în interior, vehiculele s-au apropiat, pe rând, de linia de start; arătau de parcă ar fi venit, care mai de care, ori din *Mad Max*, ori din *Curse trăsnite*. Dar aspectul nu avea nicio importanţă. Tot ce trebuiau să facă era un tur de pistă, fără nicio intervenţie umană, în mai puţin de zece ore.

Însă lucrurile nu au mers deloc ca unse! Un automobil s-a răsturnat cu roţile în sus în zona de start şi a trebuit retras din concurs.9 Motoreta abia trecuse de linia de pornire, când s-a răsturnat pe o parte şi a fost declarată eliminată din cursă. Un automobil a lovit un zid de beton după doar 45 de metri, altul s-a încurcat într-un gard din sârmă ghimpată, iar o maşinărie a fost prinsă între două tufişuri de mărăcini şi – considerându-le obiecte imobile – s-a blocat pe manevre înainte şi înapoi, înainte şi înapoi, până când cineva a trebuit, totuşi, să intervină.10 Alte vehicule s-au izbit de bolovani şi au intrat cu toată viteza în şanţuri. Osiile s-au rupt, cauciucurile s-au sfâşiat, caroseriile au zburat prin aer.11 Peisajul din jurul barului Slash X a început să semene cu un cimitir de roboţi.

Vehiculul cu cele mai bune rezultate, înscris la concurs de Universitatea Carnegie Mellon, a reuşit să parcurgă impresionanta distanţă de 11 kilometri înainte de a interpreta greşit un deal – moment în care cauciucurile au început să se rotească şi, în lipsa intervenţiei umane, au continuat să se rotească până au luat foc.12 Până la ora 11.00 dimineaţa, cursa s-a încheiat. Unul dintre organizatorii din cadrul DARPA s-a suit într-un elicopter şi a zburat spre linia de sosire, să-i informeze pe jurnaliştii care aşteptau că niciuna dintre maşinării nu avea să ajungă până acolo.13

Cursa fusese grea, plină de praf, zgomotoasă şi distructivă – şi se încheiase fără niciun câştigător. Toţi

oamenii din echipele acelea munciseră un an întreg să creeze ceva care rezistase, în cel mai bun caz, doar câteva minute. Competiţia s-a dovedit, însă, departe de a fi un dezastru. Concurenţa a dus la o explozie de noi idei, iar până la următorul Grand Challenge din 2005, tehnologia devenise aproape de nerecunoscut.

A doua oară, toţi participanţii, cu excepţia unuia singur, au depăşit distanţa de 11 kilometri atinsă în 2004. Nu mai puţin de cinci vehicule diferite au reuşit să parcurgă întreaga distanţă a cursei, de 212 kilometri, fără nicio intervenţie umană.14

Acum, cu puţin peste zece ani mai târziu, se acceptă pe scară largă ideea că viitorul aparţine mijloacelor de transport fără şofer. Pe la sfârşitul anilor ’2017, Philip Hammond, ministrul britanic de finanţe, a anunţat intenţia guvernului de a avea pe şoselele britanice, până în anul 2021, automobile fără şofer complet autonome – fără însoţitor de siguranţă la bord. Daimler a promis lansarea autovehiculelor fără şofer până în anul 2020, 15 Ford până în 202116, iar alţi fabricanţi au făcut propriile prognoze, similare cu cele citate.

Discuţiile din presă au trecut de la răsucirea pe toate părţile a întrebării dacă maşinile fără şofer urmează să apară, la dezbaterea provocărilor cărora vom fi nevoiţi să le facem faţă atunci când vor fi peste tot. „Automobilul tău fără şofer ar trebui să lovească un pieton ca să-ţi salveze ţie viaţa?”, întreba titlul unui articol din *New York Times* în iunie 2016.17 Iar în noiembrie 2017: „Ce se va întâmpla în cazul animalelor omorâte pe şosea sau al primirii unei amenzi de circulaţie, atunci când vehiculele noastre vor deţine controlul?” 18 între timp, în ianuarie 2018, *Financial Times* avertiza: „Camioanele se îndreaptă spre un viitor din care vor lipsi şoferii: sindicatele avertizează că locurile de muncă a milioane de camionagii vor fi puse în pericol”. 19

Aşadar, ce s-a schimbat? Cum s-a transformat tehnologia acestor vehicule, în numai câţiva ani, dintr-o

însăilare şubredă într-una revoluţionară şi încrezătoarea în viitorul propriu? Şi putem, oare, nutri aşteptări rezonabile ca acest progres rapid să continue?

CE ESTE ÎN JURUL MEU?

Visul nostru despre un vehicul absolut autonom datează din epoca ştiinţifico-fantastică a *jet pa*ckurilor[50], a rachetelor cosmice, a costumelor spaţiale din folie de aluminiu şi a armelor laser. În 1939, la Expoziţia Universală din New York, compania General Motors a dezvăluit viziunea sa despre viitor. Vizitatorii s-au legat de un scaun echipat cu sistem audio instalat pe o bandă transportoare care i-a purtat într-un tur de 16 minute printr-o lume imaginară.20 în spatele geamului, aceştia au văzut o machetă a visului companiei General Motors: super-autostrăzi care se întindeau de-a lungul şi de-a latul ţării, şosele care legau între ele ferme şi oraşe, străzi şi intersecţii – iar pe acestea goneau bolizi automatizaţi controlaţi prin unde radio care puteau să circule în siguranţă la viteze de până la 160 km/h. „Vi se pare cumva ciudat?" îi întreba prezentatorul. „Fantastic? Incredibil? Nu uitaţi, aceasta este lumea anilor I960!" 21

De-a lungul anilor au existat numeroase tentative de a transforma visul în realitate. General Motors a încercat acest lucru cu Firebird II, în anii 1950.22 Cercetătorii britanici s-au străduit, în anii 1960, să adapteze un Citroën DS19 astfel încât acesta să comunice cu drumul pe care rula (undeva, între Slough şi Reading, încă mai găsiţi o porţiune de vreo 14 kilometri pe care este instalat un cablu electric, rămas de la experimentele lor).23 Au urmat vehiculele din seria Navlab a celor de la Universitatea Carnegie Mellon din 1980; proiectul de un miliard de dolari şi Eureka Prometheus, al Uniunii Europene din anii 1990.24 Cu fiecare nouă încercare, visul construirii unei maşinării fără şofer părea ispititor de aproape, dar...

---

50 Aparat de zbor individual prevăzut cu dispozitive de control al zborului ce pot fi manevrate cu ambele mâini şi înzestrat cu diferite sisteme de siguranţă (n. tr.).

În aparenţă, construirea unui automobil fără şofer pare a fi relativ uşoară. Cei mai mulţi oameni reuşesc să stăpânească aptitudinile necesare pentru a conduce. Mai mult decât atât, există doar doi posibili vectori: viteza şi direcţia. E o chestiune de câtă acceleraţie să aplici şi cât de mult să răsuceşti volanul. Cât de greu poate fi?

Aşa cum a demonstrat primul Grand Challenge organizat de DARPA, construirea unui vehicul autonom este, de fapt, mult mai dificilă decât pare. Lucrurile se complică repede atunci când încerci să faci un algoritm să controleze o bucată mare şi grea de metal care merge cu 100 km/h.

Să luăm reţelele neurale, de exemplu, care sunt utilizate cu mare succes pentru depistarea tumorilor din ţesuturile mamare; ai putea crede că sunt foarte potrivite să ajute şi tehnologia unui vehicul fără şofer „să vadă” mediul înconjurător. Prin 2004, reţelele neurale (aflate încă într-o formă rudimentară faţă de versiunile avansate ale celui mai recent stadiu de dezvoltare din zilele noastre) zbârnâiau deja înăuntrul prototipurilor vehiculelor fără şofer2S, încercând să înţeleagă ceva din ceea ce transmiteau camerele de filmat instalate pe capotele automobilelor. Este cert că se poate obţine un mare volum de informaţii de valoare de la o cameră video. O reţea neurală poată să înţeleagă culoarea, textura, chiar şi caracteristicile fizice ale scenei din faţa sa – lucruri precum linii, curbe, margini şi unghiuri. Întrebarea este: ce faci cu toate aceste informaţii atunci când dispui de ele?

Desigur, i-ai putea cere automobilului: „Rulează numai pe ceva care seamănă cu asfaltul”. Dar acest lucru nu va fi de mare folos în deşert, unde şoselele nu sunt decât nişte drumuri pline de praf. I-ai putea spune: „Rulează pe cel mai neted lucru din imaginea pe care o vezi”, numai că, din păcate, cel mai neted lucru este aproape întotdeauna cerul sau o clădire cu faţada din sticlă. Ai putea să gândeşti descrierea formei drumului în termeni mai abstracţi: „Caută ceva cu două margini relativ

drepte. Liniile ar trebui să fie la o distanţă mare în partea de jos a imaginii, iar distanţa să se micşoreze treptat în partea de sus”. Pare destul de logic. Numai că, din păcate, la fel arată şi un copac într-o fotografie. În general, nu este înţelept să încurajezi un vehicul să intre într-un copac.

Problema este că aceste camere video nu-ţi pot oferi nicio idee despre proporţii sau distanţe. Este un lucru pe care regizorii de film îl folosesc în permanenţă în avantajul lor – amintiţi-vă numai scena de la începutul filmului *Star Wars* în care Star Destroyer îşi face încet-încet apariţia pe fundalul negru precum cerneala al spaţiului cosmic, planând dramatic spre partea de sus a cadrului. Ai impresia că este o namilă imensă, uriaşă, când, în realitate, a fost filmat folosind o machetă nu mai mare de un metru lungime. Este o şmecherie care dă rezultate şi arată grozav pe marele ecran. Dar, în cazul unui vehicul fără şofer pentru care două linii înguste, paralele, ar putea fi ori o şosea care continuă drept în faţă, ori trunchiul unui copac din apropiere, aprecierea corectă a distanţei devine o chestiune de viaţă şi de moarte.

Chiar dacă foloseşti mai multe camere video şi combini în mod ingenios imaginile pentru a forma o imagine 3D despre lumea din jurul tău, apare o altă posibilă problemă din cauză că te bazezi prea mult pe reţelele neurale, aşa cum a descoperit prin anii 1990 Dean Pomerleau, un om de ştiinţă de la Universitatea Carnegie Mellon. Acesta lucra la realizarea maşinăriei numite ALVINN[51], pe care o instruia cum să înţeleagă mediul înconjurător din acţiunile unui şofer uman. Pomerleau şi alţii se aşezau la volan şi îl scoteau pe ALVINN la plimbări lungi, înregistrând tot ce făceau în timp ce conduceau vehiculul. În felul acesta a fost creat setul de date de antrenament din care reţelele lor neurale trebuiau să înveţe o lecţie: „condu oriunde ar conduce şi un om, evită orice altceva”.26

51 Autonomous Land Vehicle In a Neural NetWork (engl.), „Vehicul terestru autonom într-o reţea neurală” (n. tr.).

La început, totul a funcţionat de minune. După antrenament, ALVINN a reuşit să navigheze cu uşurinţă, de unul singur, pe un drum simplu. Numai că ALVINN a întâlnit un pod şi totul a luat-o razna. Dintr-odată, a intrat într-un viraj periculos, iar Pomerleau a trebuit să apuce în grabă volanul, să prevină un accident.

După săptămâni întregi în care n-a făcut decât să analizeze informaţiile despre incident, şi-a dat seama care fusese problema: toate drumurile pe care fusese antrenat ALVINN erau mărginite de iarbă. Asemenea reţelelor neurale din capitolul „Medicina" care clasau câinii husky în funcţie de zăpada din fotografii, şi reţeaua neurală a lui ALVINN folosise iarba ca indicator principal pentru stabilirea direcţiei în care să conducă. În momentul în care iarba a dispărut, automobilul nu a mai avut habar ce să facă.

Spre deosebire de camerele video, laserele chiar *pot* să măsoare distanţele. Vehiculele care folosesc sistemul numit LIDAR (Light Detection and Ranging[52], utilizat pentru prima oară în timpul celui de-al doilea Grand Challenge organizat de DARPA în 2005) emit un foton dintr-un laser, cronometrează de cât timp are nevoie acesta pentru a se lovi de un obstacol şi a se întoarce înapoi, obţinând astfel o estimare adecvată a distanţei la care se află obstacolul respectiv. Dar nu sunt numai veşti bune: LIDAR nu ajută în niciun fel la depistarea texturii sau a culorii unui obiect, este inutil când se pune problema să interpreteze semnele de circulaţie şi nu este foarte performant pe distanţe lungi. Radarul, pe de altă parte – care face acelaşi lucru, dar folosind unde radio – se descurcă de minune, indiferent de condiţiile atmosferice, poate să detecteze obstacole aflate la mare distanţă, chiar şi să vadă prin anumite materiale, însă este absolut inutil în furnizarea vreunui detaliu despre forma sau structura

52 Numit şi radar laser, este un sistem similar radarului, care utilizează pentru emisie radiaţia laser, cu lungimi de undă în spectrul optic (inclusiv ultraviolet şi infiraroşu) (n. tr.).

obstacolului.

De sine stătătoare, niciunul dintre aceste aparate care devin surse de date – camera video, sistemul LIDAR sau radarul – nu este destul de performant pentru a înţelege ce se întâmplă în jurul unui vehicul. Pentru construirea cu succes a unui vehicul fără şofer, şmecheria este să le combini. Ceea ce ar fi o sarcină relativ uşoară dacă toate instrumentele s-ar pune de acord asupra a ceea ce văd, numai că este mult mai dificil atunci când nu fac acest lucru.

Amintiţi-vă numai de tufişul de mărăcini care a pus în încurcătură unul dintre autovehicule la primul Grand Challenge organizat de DARPA şi imaginaţi-vă că automobilul dumneavoastră fără şofer s-ar găsi într-o situaţie asemănătoare. LIDAR transmite că aveţi un obstacol în faţă. Camera video este şi ea de acord. Radarul, care poate să vadă prin tufişul firav de mărăcini, comunică însă că nu aveţi de ce să vă îngrijoraţi. Pe care dintre senzori ar trebui să-l creadă algoritmul dumneavoastră?

Ce s-ar întâmpla în cazul în care ar prevala camera video? închipuiţi-vă că vă iese în cale un autocamion mare şi alb, într-o zi noroasă. De data asta, LIDAR-ul şi radarul cad de acord că trebuie frânat, numai că pe fundalul cerului albicios şi mohorât camera video nu poate să vadă nimic care să reprezinte un pericol.

Iar dacă toate acestea nu vi se par destul de dificile, atunci aflaţi că mai există o problemă. Nu trebuie să vă faceţi griji numai din cauză că senzorii ar putea să interpreteze greşit mediul înconjurător, trebuie să ţineţi cont şi de faptul că s-ar putea să îi măsoare greşit.

Poate că aţi observat cerculeţul acela albastru din Google Maps care încercuieşte locul în care vă aflaţi; el se află acolo ca să indice o posibilă eroare în redarea semnalelor GPS. Uneori, cercul albastru va fi mic şi va marca exact poziţia în care sunteţi; alteori, el va acoperi un spaţiu mai mare şi va fi centrat într-un loc greşit. De

cele mai multe ori, asta nu contează prea mult. Ştim unde suntem şi putem să respingem informaţia inexactă. Dar un vehicul fără şofer nu are o confirmare de la faţa locului a poziţiei în care se află. Atuncâcand rulează pe o singură bandă de autostradă, mai îngustă de patru metri, nu se poate baza numai pe GPS pentru a recunoaşte exact unde se află.

GPS-ul nu este singurul care oferă coordonate ce pot fi incerte. Fiecare măsurătoare realizată de un automobil va avea o marjă de eroare oarecare: înregistrările radarului, rotaţia pneurilor, inerţia vehiculului… Nimic nu este niciodată 100% demn de încredere. Mai mult decât atât, condiţiile atmosferice înrăutăţesc lucrurile: ploaia afectează LIDAR-ul27, razele puternice ale soarelui pot să perturbe funcţionarea camerelor video, 28 iar drumurile lungi şi accidentate distrug accelerometrele.29

Aşa că, în cele din urmă, te vei alege doar cu o harababură de semnale. Întrebări care păreau simple („Unde eşti?” „Ce se află în jurul tău?” „Ce ar trebui să faci acum?”) devin uimitor de dificile. Este aproape imposibil să ştii ce să mai crezi.

Aproape imposibil. Nu chiar 100% imposibil.

Pentru că, din fericire, îţi poţi croi un drum prin tot haosul acesta – există o cale de a face supoziţii rezonabile într-o lume dezordonată, haotică. Totul se reduce la o formulă matematică extraordinar de importantă, cunoscută drept „teorema lui Bayes”.

MAREA BISERICĂ A REVERENDULUI BAYES

Nu este nicio exagerare să afirmăm că teorema lui Bayes este una dintre cele mai influente idei din istorie. Ea trezeşte un entuziasm aproape reverenţios printre oameni de ştiinţă, specialişti în învăţarea automată şi statisticieni, deopotrivă. Ideea de bază este, însă, extraordinar de simplă. Atât de incredibil de simplă, de fapt, încât, la o primă privire, aţi putea crede că nu face decât să expună ceva ce este deja evident.

Daţi-mi voie să ilustrez această idee cu un exemplu

cât de poate de banal.

Imaginaţi-vă că staţi pe un scaun şi luaţi masa la un restaurant la un moment dat, persoana care vă însoţeşte se apleacă şi vă şopteşte că a observat că Lady Gaga ia prânzul la masa vecină.

Înainte de a vă întoarce să priviţi, vă veţi întreba, desigur, câtă încredere să acordaţi spuselor prietenului dumneavoastră. Aşa că veţi lua în considerare toate cunoştinţele anterioare: calitatea restaurantului, poate, distanţa până la locuinţa artistei din Malibu sau chiar acuitatea vizuală a prietenului... Ceva de genul acesta. Dac-ar trebui să ofer o definiţie: este o convingere care poate fi cuantificată. Ca un fel de probabilitate.

Şi, în timp ce vă întoarceţi să o priviţi pe femeia în cauză, veţi folosi, în mod automat, fiecare dovadă de care dispuneţi ca să vă actualizaţi nivelul de încredere în ipoteza prietenului dumneavoastră. Poate că părul blond platinat corespunde cu ceea ce v-aţi aştepta să vedeţi la Lady Gaga, astfel încât încrederea dumneavoastră sporeşte. Însă faptul că stă la masă fără niciun bodyguard lângă ea nu mai corespunde, astfel încât încrederea scade. Ideea este că fiecare detaliu nou contribuie cu ceva la evaluarea generală pe care o întreprindeţi.

Asta este tot ce face teorema lui Bayes: oferă un mod sistematic de a ne actualiza nivelul de încredere într-o ipoteză pe baza unor dovezi.30 Ea acceptă faptul că nu poţi fi niciodată absolut sigur de teoria la care te gândeşti, dar îţi permite să faci supoziţia cea mai bună pe baza informaţiilor de care dispui. Aşadar, când vă daţi seama că femeia de la masa vecină poartă o rochie confecţionată din came[53] – o alegere stilistică pe care este puţin probabil să o întâlniţi din întâmplare la vreo fiinţă care să nu fie Lady Gaga acest lucru ar putea fi suficient pentru ca încrederea dumneavoastră să depăşească pragul şi să trageţi

53 Lady Gaga este artista care, în 2010, a apărut la gala de decernare a premiilor Video Music Awards îmbrăcată într-o rochie confecţionată din came (n. tr.).

concluzia că în acel restaurant se află, într-adevăr, Lady Gaga.

Însă teorema lui Bayes nu este numai o ecuaţie care defineşte modul în care oamenii iau decizii. Este mult mai importantă de atât. Pentru a o cita pe Sharon Bertsch McGrayne, autoarea cărţii *The Theory That Would Not Die*:

[Teorema lui] Bayes vine în contradicţie cu profunda noastră convingere că ştiinţa modernă are nevoie de obiectivitate şi de precizie.31

Oferind un mecanism de măsurare a încrederii pe care o ai într-un anumit lucru, teorema lui Bayes îţi permite să tragi concluzii de bun-simţ din observaţii sumare, din informaţii dezordonate, incomplete şi aproximative – ba chiar şi din neştiinţa ta.

Teorema lui Bayes nu există numai ca să confirme intuiţiile existente. Se dovedeşte că, dacă eşti nevoit să-ţi cuantifici convingerile într-un anumit lucru, se ajunge adesea la concluzii paradoxale, care par să contrazică intuiţia. Teorema lui Bayes explică de ce mai mulţi bărbaţi decât femei sunt identificaţi în mod greşit drept viitori ucigaşi (vezi exemplul din capitolul „Justiţia”). Şi tot teorema lui Bayes explică de ce – chiar dacă ai fost diagnosticată ca suferind de cancer mamar –, dat fiind nivelul semnificativ de eroare al testelor, înseamnă că probabil nu ai aşa ceva (vezi capitolul „Medicina”). Teorema lui Bayes este, în toate domeniile ştiinţei, un instrument puternic pentru selectarea şi înţelegerea a ceea ce cunoaştem de fapt.

Însă modul de gândire bayesian îşi arată cu adevărat valoarea atunci când încerci să iei simultan în considerare mai mult de o singură ipoteză – de exemplu, atunci când te străduieşti să îţi dai seama ce nu este în regulă cu un pacient pe baza simptomelor pe care le manifestă[54] sau să

54 Watson, calculatorul IBM despre care am discutat în capitolul „Medicina”, foloseşte din plin aşa-numita „inferenţă bayesiană”. Vezi postarea „Watson and Healthcare”, on-line la adresa: www.ibm.com/developerworks/library/os-ind-watson/ (accesată cel mai recent la 19 martie 2019) (n. a.).

găseşti locul unde se află un vehicul fără şofer pe baza indicaţiilor oferite de senzorii acestuia. Teoretic, orice boală sau orice punct de pe o hartă ar putea să reprezinte un adevăr întemeiat, bazat pe date observate. Tot ce ai de făcut este să evaluezi elementele ca să decizi care are şansele cele mai mari să fie corect.

Iar din acest moment, găsirea poziţiei unui automobil fără şofer se dovedeşte a fi destul de asemănătoare cu ceea ce l-a pus în dificultate pe Thomas Bayes, pastorul prezbiterian şi talentatul matematician britanic al cărui nume îl poartă teorema. Pe la mijlocul anilor 1700, Bayes a scris un eseu în care a inclus detaliile unui joc pe care îl inventase pentru explicarea problemei. Era ceva de genul acesta: 32

Imaginaţi-vă că staţi cu spatele la o masă pătrată. Fără să vedeţi, eu arunc pe masă o minge roşie.

Dumneavoastră trebuie să ghiciţi unde a aterizat aceasta. Nu va fi uşor: fără nicio informaţie de la care să puteţi pleca, nu există nicio modalitate reală de a şti unde ar putea fi mingea pe masă.

Atunci, ca să vă ajut să ghiciţi, arunc o a doua minge, de culoare diferită, pe aceeaşi masă.

Trebuie să faceţi acelaşi lucru, şi anume să stabiliţi unde se află prima minge, cea roşie, însă acum vă voi spune unde ajunge pe masă cea de-a doua minge faţă de prima: dacă este în faţa ei, în spatele ei, la stânga sau la dreapta mingii roşii. Iar dumneavoastră vă veţi putea corecta evaluarea iniţială.

Apoi vom repeta: voi arunca a treia, a patra şi a cincea minge pe masă şi de fiecare dată vă voi spune unde aterizează fiecare dintre ele în raport cu prima minge, cea roşie – a cărei poziţie pe masă încercaţi să o ghiciţi.

Cu cât voi arunca mai multe mingi şi cu cât vă voi oferi mai multe informaţii, cu atât imaginea poziţiei mingii roşii ar trebui să vă fie tot mai clară.

Nu veţi fi niciodată absolut sigur unde se află cu exactitate, dar vă puteţi actualiza nivelul de încredere cu

privire la poziţia ei până când veţi sfârşi prin a găsi un răspuns care să vă satisfacă pe deplin.

Într-un fel, poziţia reală a vehiculului fără şofer este asemănătoare cu aceea a mingii roşii.

În locul unei persoane care stă cu spatele la masă, avem de-a face cu un algoritm care încearcă să evalueze cu precizie unde se află vehiculul la un moment dat, iar în locul celorlalte mingii aruncate pe masă, avem sursele de informaţii: GPS-ul, măsurarea inerţiei şi aşa mai departe. Niciuna dintre acestea nu-i spune algoritmului unde se află automobilul cu exactitate, dar fiecare adaugă încă puţină informaţie, pe care algoritmul o poate folosi pentru a-şi actualiza nivelul de încredere.

Este o şmecherie cunoscută drept „inferenţă probabilistică” – folosirea informaţiilor (plus teorema lui Bayes) pentru a face o inferenţă despre poziţia reală a obiectului. Şi, „ambalată” în mod corespunzător, devine doar încă unul dintre algoritmii de învăţare automată.

Pe la începutul acestui mileniu, inginerii acumulaseră deja suficientă experienţă practică cu rachetele de croazieră, navetele spaţiale şi aparatele de zbor, ca să ştie cum să abordeze problema poziţiei. A face un vehicul fără şofer să răspundă la întrebarea „Unde mă aflu?” nu era încă o însărcinare banală, dar, cu puţină gândire bayesiană, măcar era un lucru realizabil.

De la adevăratul cimitir de roboţi de la Grand Challenge din 2004, la uimitorul triumf al tehnologiei de la evenimentul din 2005 – în timpul căruia cinci vehicule diferite au reuşit să parcurgă peste 160 de kilometri fără nicio intervenţie umană – multe dintre cele mai mari salturi înainte au avut loc datorită lui Bayes. Algoritmii care se bazau pe ideile bayesiene au ajutat la soluţionarea celorlalte întrebări la care mai trebuia să răspundă un autovehicul fără şofer: „Ce se află în jurul meu?” şi „Ce ar trebui să fac acum?”[55]

55 Câştigătoarea cursei din 2005, care fusese creată de o echipă de la Universitatea Stanford, a fost descrisă destul de bine de

AŞADAR, AUTOMOBILUL TĂU FĂRĂ ŞOFER AR TREBUI

SĂ LOVEASCĂ UN PIETON CA SĂ-ŢI SALVEZE VIAŢA?

Să ne oprim o clipă şi să ne gândim la cea de-a doua întrebare. Pentru că în anul 2016, pe la începutul toamnei, retras într-un colţ mai liniştit al sălii de expoziţie, altfel plină de agitaţie, a Salonului Auto de la Paris, un purtător de cuvânt al companiei Mercedes-Benz a făcut un anunţ destul de ieşit din comun pe această temă. Christoph von Hugo, directorul departamentului de sisteme de asistenţă la volan şi sisteme de siguranţă al companiei, a fost întrebat, într-un interviu, ce ar trebui să facă un vehicul fără şofer în cazul unui accident. Iar el a răspuns:

Dacă ştii că poţi salva cel puţin o persoană, atunci salveaz-o măcar pe aceea.33

Este doar o chestiune de logică pură, aţi putea crede. În niciun caz nu merită să devină subiectul unei ştiri de pe prima pagină a ziarelor.

Numai că Hugo nu fusese întrebat despre cine ştie ce accident de demult. El fusese pus la încercare şi rugat să ofere un răspuns la un bine-cunoscut experiment de gândire, conceput în anii 1960, care descria un tip aparte de coliziune. Reporterul se referea la o dilemă morală ciudată, care te forţează să alegi între două rele, cunoscută drept „problema vagonului”, după tramvaiul scăpat de sub control descris în formularea iniţială. În cazul automobilelor autonome, ea sună astfel:

Imaginaţi-vă că, peste câţiva ani, vă aflaţi într-un vehicul autonom şi conduceţi fără nicio grijă pe o stradă dintr-un oraş. În faţa dumneavoastră, un semafor indică roşu, însă o defecţiune mecanică a automobilului vă pune în imposibilitatea de a-l opri. O coliziune este inevitabilă, însă vehiculul poate să facă o alegere: să vireze, părăsind drumul, şi să intre într-un zid de beton, ceea ce va duce la

matematicianul Pesri Diaconis, de la aceeaşi universitate, care a declarat: „Fiecare şurub al vehiculului aceluia era bayesian” (n. a.).

moartea sigură a tuturor celor din vehicul. Sau să-şi continue traiectoria, salvând vieţile tuturor celor din interior, dar să omoare pietonii care traversează strada. Cum ar trebui programat vehiculul să acţioneze? Cum decideţi cine ar trebui să moară?

Fără îndoială că aveţi propria opinie. Poate credeţi că maşina ar trebui să încerce să salveze cât mai multe vieţi cu putinţă. Sau poate consideraţi că porunca biblică „să nu ucizi" ar trebui să prevaleze asupra oricăror altor considerente, lăsându-l pe cel aflat în maşină să suporte consecinţele.[56]

Hugo a fost lipsit de echivoc în privinţa poziţiei companiei Mercedes. „Trebuie salvat cel din maşină!" Şi a continuat: „Dacă ştii cu siguranţă că o moarte poate fi prevenită, atunci aceea este prioritatea."[57]

În zilele care au urmat interviului, internetul a fost inundat de articole care incriminau atitudinea companiei. „Maşinile lor vor acţiona foarte asemănător cu modul în care ar face-o un răzgâiat şofer european al unei maşini de lux" 34, scria autorul unui articol. Într-adevăr, într-un sondaj publicat în revista *Science* în aceeaşi vară, 35 76% dintre respondenţi simţeau că ar fi un gest mai etic din partea vehiculelor autonome ca acestea să salveze cât mai multe vieţi cu putinţă, omorând astfel persoanele din interiorul lor. Mercedes nimerise de partea opusă a opiniei generale.

Aşa să fie, oare? Pentru că atunci când, în acelaşi sondaj, participanţii au fost întrebaţi dacă ei înşişi chiar *ar cumpăra o* maşină care i-ar ucide dacă împrejurările ar

---

56 în presă, de la New York Times la Mail on Sunday, au apărut numeroase versiuni ale acestui scenariu. „Dacă pietonul este o bunicuţă de 90 de ani? Dacă este un copilaş? Dacă în automobil s-ar afla un laureat al Premiului Nobel?" Toate au la bază aceeaşi dilemă (n. a.).

57 Citatele din declaraţia lui Cristoph von Hugo au fost preluate din articolul postat on-line la adresa: http://www.automarket.ro/stiri/mercedes-a-rezolvat-problema-morala-masinile-autonome- vor-73753.html (accesată cel mai recent la 15 martie 2019) (n. tr.).

cere-o, aceştia au devenit dintr-odată reticenţi în a se sacrifica pe ei înşişi pentru binele general.

Aceasta este o dilemă morală care face ca opiniile să fie împărţite – şi nu numai în privinţa răspunsului pe care oamenii îl consideră cel mai nimerit. Ca experiment de gândire, rămâne favoritul de necontestat al reporterilor specializaţi în tehnologie şi al altor jurnalişti, însă toţi specialiştii în automobile autonome pe care i-am intervievat au dat ochii peste cap imediat ce am amintit de problema vagonului. Eu, personal, încă mai am o slăbiciune faţă de ea. Simplitatea sa ne sileşte să recunoaştem ceva important despre vehiculele fără şofer, să reflectăm la ceea ce simţim faţă de un algoritm care face o judecată de valoare în privinţa vieţii noastre şi a altor persoane. La temelia acestei noi tehnologii – aşa cum este cazul cu aproape toţi algoritmii – se află aspecte care privesc puterea, aşteptările, controlul şi delegarea responsabilităţii. Precum şi dacă ne putem aştepta ca tehnologia pe care o inventăm să ni se potrivească nouă şi nu viceversa. Înţeleg însă şi reacţia rezervată pe care o primesc din partea celor care se ocupă de crearea automobilelor autonome. Ei ştiu, mai bine decât oricine altcineva, cât de departe suntem de momentul în care chiar va trebui să ne facem griji din cauza problemei vagonului.

ÎNCĂLCAREA REGULILOR DE CIRCULAŢIE

Teorema lui Bayes şi puterea probabilităţii au propulsat o mare parte a inovaţiei în domeniul automobilelor autonome încă de la prima competiţie organizată de DARPA. L-am întrebat pe Paul Newman, profesor de robotică la Universitatea Oxford şi fondatorul Oxbotica, o companie care construieşte vehicule fără şofer şi le testează pe şoselele din Marea Britanie, cum se comportau cele mai recente automobile autonome ale sale, iar el mi-a explicat următoarele:

Sunt multe, multe milioane de linii de cod, dar am reuşit să încadrez toată chestia drept o inferenţă

probabilistică. Pe toată! 36

Dar în timp ce inferenţa bayesiană se îndreaptă oarecum spre explicarea modului în care este posibilă existenţa maşinilor fără şofer, aceasta arată totodată şi că autonomia deplină, lipsită de orice intervenţie a unui şofer uman, este o problemă foarte, foarte dificil de rezolvat.

Închipuiţi-vă ceea ce sugerează Paul Newman: „ai două vehicule care se apropie cu viteză unul de celălalt” – rulând, să spunem, în direcţii opuse pe o şosea A[58] în curbă lină. Un şofer uman se va simţi confortabil în acest scenariu, ştiind că maşina cealaltă va rămâne pe banda ei de mers şi că va trece de el în siguranţă, la vreo doi metri de marginea marcajului. „Însă, pentru câteva minute – explică Newman – totul arată ca şi cum maşinile vor intra în coliziune”. Cum înveţi un automobil autonom să nu intre în panică într-o asemenea situaţie? Nu-ţi doreşti ca vehiculul să devieze de pe şosea, încercând să evite o coliziune care nu va avea loc niciodată. De asemenea, nu ţi-ai dori nici să rămână nepăsător dacă într-adevăr te găseşti în pragul unei coliziuni frontale. Aduceţi-vă aminte că aceste maşini nu fac decât să aprecieze cu aproximaţie ce au de făcut. Cum să le facem să ghicească de fiecare dată corect? Asta, spune Newman, „este o problemă foarte, foarte dificilă”.

Este o problemă care i-a pus mult timp în impas pe specialişti, însă ea are o soluţie. Trucul este să incluzi din construcţie un model care să ştie cum se vor comporta ceilalţi şoferi (cei care sunt sănătoşi la minte). Din nefericire, nu se poate spune acelaşi lucru şi despre alte scenarii de conducere particulare.

Newman explică: „Cea mai dificilă parte o constituie problemele care apar în timp ce şofezi şi care nu au nimic de-a face cu conducerea autovehiculului”. De exemplu, să înveţi un algoritm să înţeleagă că trebuie să fie extrem de

58 în Marea Britanie, o şosea A este o cale de circulaţie principală, mai îngustă decât o autostradă, dar mai lată şi mai dreaptă decât o şosea B (n. tr.).

atent atunci când aude melodia unei maşini din care se vinde îngheţată sau când trece pe lângă un grup de copii care se joacă liniştiţi cu mingea pe trotuar. Sau să recunoască ţopăitul derutant al unui cangur – o chestiune pe care, în momentul în care am scris această carte, compania Volvo a recunoscut că se chinuia să o soluţioneze.37 Probabil că nu este o problemă chiar atât de importantă în rusticul Surrey, dar este ceva ce maşinile trebuie să-şi însuşească pentru a corespunde cerinţelor de pe drumurile publice din Australia.

Iată ceva încă şi mai greu de realizat: cum înveţi o maşină că, uneori, ar trebui să încalce regulile de circulaţie? Ce ar fi dacă staţionezi la semafor şi cineva apare brusc în faţa maşinii şi îţi face frenetic semn să avansezi? Sau dacă o ambulanţă cu semnalele luminoase aprinse încearcă să străbată o stradă îngustă, iar tu trebuie să te urci pe trotuar ca s-o laşi să treacă? Sau dacă o cisternă care transportă petrol s-a răsturnat pe un drum de ţară, iar tu trebuie să părăseşti locul cu orice preţ?

„Nicio situaţie de acest gen nu figurează în Codul rutier”, subliniază Newman în mod corect. Şi, totuşi, o maşină cu adevărat autonomă trebuie să ştie cum să facă faţă tuturor acestor împrejurări, dacă dorim ca aceasta să funcţioneze fără *absolut* nicio intervenţie umană. Chiar şi în situaţii de urgenţă.

Ceea ce nu înseamnă că aceste probleme sunt insurmontabile. „Nu cred că există vreun nivel de inteligenţă la care să nu reuşim să aducem o maşină”, mi-a spus Newman. „Singura întrebare care se pune este când vom reuşi acest lucru”.

Din păcate, răspunsul la dilema de mai sus este: probabil că nu în viitorul apropiat. Visul realizării unui automobil autonom, care ne dorim cu toţii să se îndeplinească, ar putea fi cu mult mai îndepărtat decât credem.

Asta, pentru că mai există un nivel de dificultate pe care trebuie să-l depăşim atunci când încercăm să

construim această fantasmagorie ştiinţifico-fantastică de vehicul fără şofer, fără volan, care te duce oriunde şi face orice. Iar el depăşeşte cu mult provocările tehnologice. Un automobil complet autonom trebuie să se descurce şi cu problema spinoasă reprezentată de factorul uman.

Jack Stil goe, sociolog la University College London şi specialist în impactul social al tehnologiei, explică:

Oamenii sunt răutăcioşi. Ei sunt agenţi activi, nu doar elemente pasive într-un scenariu.38

Imaginaţi-vă, pentru o clipă, o lume în care există vehicule cu adevărat şi perfect autonome. Prima regulă a algoritmilor de la bord va fi să evite coliziunile ori de câte ori este posibil. Iar aceasta schimbă relaţiile dintre participanţii la trafic de pe drumuri. Dacă stai în faţa unui automobil fără şofer, acesta trebuie să oprească. Dacă intri în faţa unui vehicul autonom la o intersecţie, acesta trebuie să-ţi cedeze trecerea.

Cum a declarat unul dintre participanţii la un *focus group* organizat în 2016 la London School of Economics: „Pur şi simplu, le vei lua faţa. Se vor opri, iar tu vei trece pe lângă ele”. În traducere: aceste vehicule vor putea fi intimidate.

Stil goe este de aceeaşi părere:

Persoanele care până acum au fost oarecum lipsite de putere pe şosele, cum ar fi bicicliştii, ar putea începe să pedaleze foarte încet în faţa automobilelor autonome, ştiind că acestea nu vor putea face niciodată niciun gest agresiv.

Soluţionarea problemei ar putea însemna adoptarea unor reguli mai stricte împotriva acelora care abuzează de condiţia lor de ciclişti sau de pietoni. Este o măsură care a mai fost luată şi altă dată, bineînţeles: gândiţi-vă numai la interzicerea traversării prin locuri nepermise. Sau ar putea însemna eliminarea forţată a orice altceva mai circulă pe şosele – aşa cum s-a întâmplat în momentul apariţiei autoturismelor –, motiv pentru care, în prezent, nu mai întâlnim biciclete, cai, care, căruţe sau pietoni pe o

autostradă.

Dacă ne dorim vehicule complet autonome, atunci, aproape cu certitudine, vom fi nevoiţi să luăm măsuri asemănătoare şi pentru a limita numărul şoferilor agresivi, al maşinilor de îngheţată, al copiilor care se mai joacă încă în stradă, al semnelor rutiere care anunţă „lucrări pe partea carosabilă”, dar şi numărul pietonilor dificili, al ambulanţelor, al bicicliştilor, al scuterelor şi al oricui altcuiva care face problema autonomiei atât de dificilă. Toate ar fi bune şi frumoase, numai că viziunea aceasta diferă puţin de modul în care ne sunt prezentate nouă lucrurile în prezent, pentru că, aşa cum mi-a declarat Stil go e:

Toate cuvântările frumoase despre autonomie şi mijloacele de transport au de-a face, mai degrabă, cu modul în care să *nu* schimbăm lumea. Cu alte cuvinte, cum să păstrăm lumea aşa cum este, dar să construim şi să permitem unui robot să fie tot atât de priceput ca o fiinţă umană, apoi încă şi mai priceput decât aceasta la conducerea unei maşini. Iar eu cred că asta este o mare prostie.

„Dar ia stai puţin!”, ar putea gândi unii dintre voi. „Problema asta nu a fost rezolvată deja? Waymo, automobilul autonom al lui Google, nu a condus deja milioane de kilometri? Waymo-urile, maşinile complet autonome (sau cel puţin *aproape* complet autonome) nu circulă în prezent pe străzile din Phoenix, statul Arizona?”

Ei bine, da! Aşa este! Dar nu orice kilometru de şosea este la fel. Cele mai multe distanţe sunt uşor de parcurs, poţi conduce şi dacă eşti cu gândul aiurea. Altele sunt mult mai dificile. În momentul în care scriu această carte, automobilelor Waymo nu li se permite să meargă chiar pe oriunde: sunt „limitate geografic” la o suprafaţă redusă şi predefinită. Tot astfel sunt şi vehiculele fără şofer pe care companiile Daimler şi Ford îşi propun să le lanseze pe şosele până în anul 2020, respectiv 2021. Sunt maşini care vin doar la comandă, ducând mai mulţi pasageri care merg

(aproape) în aceeaşi direcţie, iar călătoriile le sunt limitate la o zonă de circulare predeterminată. Iar asta simplifică destul de mult problema autonomiei.

Paul Newman crede că acesta este viitorul la care ne putem aştepta din partea vehiculelor fără şofer:

Ele vor ieşi să lucreze într-o zonă pe care o cunosc foarte bine, unde posesorii lor sunt foarte convinşi că vor funcţiona. Prin urmare, ar putea fi o anumită parte a oraşului, nu în mijlocul unei zone cu şosele mai neobişnuite sau unde s-ar putea să le iasă nişte vaci în cale. Poate vor funcţiona în anumite momente ale zilei sau în anumite condiţii atmosferice. Vor fi exploatate ca un serviciu de transport în comun.

Nu este chiar acelaşi lucru cu autonomia deplină. Iată, în cele ce urmează, opinia lui Jack Stil goe despre compromisul necesar:

Ceea ce seamănă cu un sistem autonom este, de fapt, un sistem pe care lumea este constrânsă să-l facă să pară autonom.

Viziunea în care am ajuns să credem este ca o iluzie optică. Un miraj care promite o limuzină cu un şofer particular stilat pentru toată lumea, însă care, dacă o cercetăm mai bine, nu va fi decât un microbuz local.

Dacă încă mai aveţi nevoie de argumente, îi voi lăsa ultimul cuvânt uneia dintre cele mai importante reviste auto din America, şi anume *Car and Driver*:

Nicio companie producătoare de maşini nu se aşteaptă ca utopia futuristă a unor străzi lipsite de accidente şi înţesate de vehicule fără şofer să prindă contur în viitorul apropiat şi nici mai devreme de câteva decenii. Companiile vor, însă, să fie luate în serios de bursa de pe Wall Street, dar şi să înflăcăreze imaginaţia unui public din ce în ce mai puţin interesat de condusul unui automobil. Iar, între timp, speră să vândă o mulţime de autovehicule dotate cu cea mai recentă şi mai sofisticată tehnologie de asistare la conducere (*driverassistance technology*).*39*

Dar ce este cu tehnologia asta de asistare la conducere? în definitiv, maşinile fără şofer nu sunt o propunere de genul totul sau nimic.

În categoria tehnologiei „fără şofer" există şase niveluri diferite: de la nivelul 0 – fără niciun fel de automatizare şi până la nivelul 5 – visul autonomiei complete. Între acestea, ea variază de la *cruise control* (sau regulator de viteză) (nivelul 2), la vehiculele autonome cu delimitare zonală (nivelul 4), iar în limbajul curent sunt cunoscute drept: nivelul 1 (*feet ofl*), nivelul 2 (*hands off*), nivelul 3 (eyes *ofl*) şi nivelul 4 (*brain ofl*).[59]

Aşa că nivelul 5 poate că nu este chiar în orizontul nostru imediat de aşteptare, iar nivelul 4 nu va fi chiar ceea ce pare, însă ne mai aşteaptă foarte multă automatizare în timp ce progresăm la nivelul superior. Oare chiar ne-ar deranja să trecem încet-încet de la un nivel la altul cu automobilele noastre personale? Şi să construim maşini cu volane, pedale de frână, şoferi adevăraţi pe scaunul şoferului, lăsând fiinţa umană să intervină şi să preia comanda doar în situaţiile de urgenţă? Chiar nu ne vom mulţumi cu aceste progrese până când va avansa suficient tehnologia?

Din nefericire, lucrurile nu sunt chiar atât de simple. Pentru că povestea aceasta mai are o întorsătură neaşteptată. O pleiadă întreagă de alte probleme, de fapt.

---

59 în conformitate cu definiţiile propuse de Society of Automotive Engineers: nivelul 1 (feet off, hands on), şoferul împarte controlul vehiculului cu algoritmul (ca în cazul parcării asistate sau al monitorizării automate a vitezei); nivelul 2 (hands off) presupune că sistemul automat preia pe deplin controlul vehiculului, accelerând, punând frână şi răsucind volanul, în vreme ce şoferul trebuie să intervină doar în cazul în care sistemul nu face faţă; nivelul 3 (eyes off) înseamnă că automobilul se poate descurca singur chiar şi în cazurile în care ar fi nevoie să frâneze brusc, ceea ce înseamnă că şoferul uman îşi poate îndrepta atenţia spre alte sarcini (cum ar fi să scrie un SMS), deşi trebuie să fie pregătit să intervină în cazuri extreme. Nivelul 4 de automatizare (brain/mind off) seamănă cu nivelul 3, numai că şoferul mai trebuie să intervină doar în cazuri cu totul speciale (într-un ambuteiaj etc.), vehiculul fiind capabil să ia chiar şi decizii legate de

Un obstacol inevitabil pentru orice altceva în afară de conducerea complet lipsită de elementul uman.

MEZINUL COMPANIEI

Printre piloţii Air France, Pierre-Cădric Bonin era cunoscut drept „mezinul companiei”.40 El s-a alăturat Air France la frageda vârstă de 26 de ani, având la activ doar câteva sute de ore de zbor, şi a crescut profesional printre flota de Airbus-uri a companiei. Până când a urcat la bordul fatidicului AF447, la vârsta de 32 de ani, reuşise să acumuleze respectabilul număr de 2 936 de ore de zbor, deşi asta îl făcea să fie totuşi cel mai puţin experimentat dintre cei trei piloţi aflaţi în aeronavă.41

Cu toate acestea, Bonin a fost cel care s-a aşezat la comanda zborului 447 al Air France pe 31 mai 2009, a scos avionul de pe pista Aeroportului International Galcão din Rio de Janeiro şi s-a îndreptat spre casă, spre Paris.42

Era un Airbus A330, unul dintre cele mai sofisticate avioane comerciale construite vreodată. Sistemul său autopilot era atât de modern, încât putea, practic, să efectueze un întreg zbor neasistat, cu excepţia manevrelor de decolare şi aterizare. Chiar şi atunci când pilotul era la comandă, avionul avea încorporate diverse mecanisme de siguranţă pentru a reduce la minimum riscul de apariţie al erorilor umane.

Numai că întotdeauna există un pericol nebănuit care apare dacă realizezi un sistem automat care poate gestiona în siguranţă aproape orice problemă pe care au reuşit s-o anticipeze creatorii săi: dacă se aşteaptă de la un pilot să preia manevrarea avionului numai în împrejurări excepţionale, acesta nu-şi va mai păstra aptitudinile necesare pentru a opera singur sistemul, pentru a conduce neajutat aeronava. Aşa că omul va ajunge să aibă doar o experienţă foarte redusă la care să se raporteze atunci când va întâlni o situaţie de urgenţă pe care nimeni n-a prevăzut-o.

Şi exact asta s-a întâmplat cu zborul Air France 447: cu toate că Bonin acumulase mii de ore de antrenament în

carlinga unui Airbus, experienţa lui reală de pilotare manuală a unui A330 era minimă. Rolul său ca pilot fusese, în cea mai mare parte, acela de a monitoriza sistemul automat. Asta însemna că atunci când pilotul automat s-a deconectat, în timpul acelui zbor de noapte, Bonin nu a ştiut cum să piloteze avionul în siguranţă.43

Dificultăţile au început când s-au format cristale de gheaţă în interiorul senzorilor din fuzelaj care măsurau viteza vântului. Incapabil să facă o măsurătoare precisă, pilotul automat a declanşat o alarmă în cabina piloţilor şi a transferat responsabilitatea echipajului uman. Aceasta nu era o chestiune îngrijorătoare în sine. Însă atunci când avionul a ajuns într-o zonă restrânsă cu turbulenţe, Bonin cel lipsit de experienţă a reacţionat exagerat: când avionul a început să se răsucească uşor pe partea dreaptă, Bonin a apucat manşa şi a tras-o spre stânga, în acelaşi moment, ceea ce este foarte important, a tras manşa înapoi, ceea ce a făcut ca avionul să urce brusc şi abrupt.44

Pe când aerul din jurul avionului se rarefia, Bonin continua să tragă cu putere de manşă înapoi, până ce botul avionului a ajuns atât de înălţat, încât aerul nu mai putea să circule lin de-a lungul aripilor, aşa că acestea s-au transformat în adevărate paravânturi, iar aeronava şi-a pierdut brusc portanţa, intrând în cădere liberă cu botul îndreptat în sus.

Alarmele sunau în carlingă. Căpitanul a năvălit înăuntru din cabina de odihnă. AF447 se prăbuşea în ocean cu o viteză de 3.000 de metri pe minut.

De acum, cristalele de gheaţă se topiseră, nu mai exista nicio defecţiune mecanică, iar oceanul era destul de departe sub ei pentru ca piloţii să poată redresa avionul la timp. Bonin şi copilotul său ar fi putut să-i salveze cu uşurinţă pe toţi cei de la bord, în numai 10 – 15 secunde, prin simpla împingere a manşei înainte, aplecând botul avionului în jos şi lăsând aerul să circule din nou lin peste aripile aparatului de zbor.45

Din cauza panicii care îl cuprinsese, Bonin lăsase

însă manşa trasă înapoi. Nimeni nu şi-a dat seama că el provocase toată drama. S-au scurs secunde preţioase. Căpitanul a sugerat să echilibreze aripile avionului. S-au întrebat, repede, dacă avionul cobora sau urca. Atunci, la 2 400 de metri distanţă de nivelul mării, copilotul a preluat controlul avionului.46

„Urcă… urcă… urcă… urcă…”, se aude copilotul strigând.

„Dar am avut tot timpul manşa trasă înapoi!” răspunde Bonin.

Atunci căpitanului i-a picat fisa. Şi-a dat, în sfârşit, seama că de mai mult de trei minute erau în cădere liberă, în stare de blocaj aerodinamic şi le-a ordonat să coboare botul avionului. Prea târziu, însă. În mod tragic, ajunseseră prea aproape de suprafaţa apei. Bonin strigă: „Fir-ar să fie! Ne vom prăbuşi. Nu se poate!” 47 Câteva clipe mai târziu, aeronava a plonjat în Oceanul Atlantic, omorându-i pe toţi cei 228 de oameni de la bord.

Cu 26 de ani înainte de prăbuşirea avionului Air France, în 1983, psihologul Lisanne Bainbridge a scris o lucrare fundamentală despre pericolele care ne pândesc atunci când ajungem să ne bazăm prea mult pe sistemele automate.48 Construieşte o maşină care îmbunătăţeşte performanţele umane, arăta ea, şi aceasta va duce – în mod ironic – la o diminuare a aptitudinilor umane.

În prezent, suntem cu toţii, într-o oarecare măsură, martorii acestui fenomen. De aceea oamenii nu-şi mai pot aminti numerele de telefon, de aceea mulţi dintre noi ne chinuim să ne descifrăm scrisul de mână şi de aceea foarte mulţi nu mai pot să ajungă nicăieri fără GPS. Dată fiind tehnologia, care face totul pentru noi, mai rămân puţine prilejuri să ne exersăm abilităţile.

Există o anumită îngrijorare că acelaşi fenomen s-ar putea să se manifeste şi odată cu apariţia vehiculelor care se vor conduce singure – unde mizele vor fi cu mult mai mari decât în cazul scrisului de mână. Până când vom ajunge la autonomia deplină, maşina va mai preda uneori

controlul, pe neaşteptate, şoferului. Vom reuşi să ne aducem instinctiv aminte ce avem de făcut? Iar şoferii adolescenţi din viitor vor avea vreodată şansa să ajungă să stăpânească toate aptitudinile necesare?

Chiar dacă toţi şoferii reuşesc să rămână competenţi [60] (şi interpretăm cu multă generozitate verbul „să rămână”), mai există o problemă cu care va trebui să ne confruntăm. Pentru că tot ce i se cere şoferului uman să facă înainte de întreruperea pilotului automat este, de asemenea, important. Există numai două posibilităţi. Şi – aşa cum subliniază Bainbridge – niciuna nu este foarte atrăgătoare.

O maşină de nivelul 2, *hands-off*, se va aştepta ca şoferul să fie foarte atent la drum în orice moment.49 Nu este suficient de performanţă, încât să fie lăsată de una singură şi va avea nevoie de o supraveghere atentă. Revista *Wired* descria la un moment dat acest nivel astfel: e „ca atunci când îi ceri unui copilaş care abia a învăţat să meargă să te ajute să speli vasele”. 50

În momentul scrierii cărţii de faţă, sistemul autopilot al unei maşini Tesla este un exemplu în acest sens.51 în prezent, este ca o funcţie mai sofisticată de *emise control* – va conduce, va frâna şi va accelera pe autostradă, dar se aşteaptă ca şoferul să fie vigilent şi atent, pregătit să intervină în orice moment. Ca să se asigure că eşti atent, ea va declanşa o alarmă dacă îţi iei mâinile de pe volan prea mult timp.

Dar, aşa cum a scris Bainbridge în lucrarea ei, această abordare va ajunge să provoace accidente, pentru că este, pur şi simplu, lipsit de realism să te aştepţi ca

---

60 Există diverse măsuri care se pot lua pentru a remedia problemele care apar din cauză că nu ai avut prea mult timp să exersezi conducerea manuală a unui avion. De exemplu, de la prăbuşirea avionului Air France se pune acum accentul pe antrenarea noilor piloţi, astfel încât să fie capabili să piloteze avionul atunci când pilotul automat nu reuşeşte să o facă. De asemenea, ei sunt stimulaţi ca, la intervale regulate, să oprească sistemul automat pentru a-şi păstra îndemânarea de a pilota manual avionul (n. a.).

oamenii să fie vigilenţi:

Este imposibil, chiar şi pentru o persoană foarte motivată, să-şi păstreze atenţia vizuală îndreptată în mod eficient spre o sursă de informaţii în care se întâmplă foarte puţine lucruri, pentru mai mult de aproximativ o jumătate de oră.52

Sunt unele dovezi că oamenii au întâmpinat dificultăţi în a ţine cont de insistenţa companiei Tesla ca ei să fie atenţi la drum. Joshua Brown, care a murit la volanul maşinii sale Tesla în anul 2016, utiliza funcţia de autopilot de 37 de minute şi 30 de secunde când maşina a lovit un camion care traversa banda pe care conducea el. Din cercetările efectuate de Administraţia pentru Siguranţa Traficului Rutier a rezultat că Brown nu se uita la drum în momentul impactului.53 Accidentul a ajuns pe prima pagină a ziarelor din toată lumea, însă asta nu i-a oprit pe unii nesăbuiţi de pe YouTube să posteze materiale video în care se arăta, cu mult entuziasm, cum să-ţi păcăleşti maşina, astfel încât aceasta să creadă că eşti atent la drum. Chipurile, dacă lipeşti o doză de Red Bull de volan 54 sau îndeşi o portocală în volan55, asta va împiedica maşina să declanşeze alarmele acelea sâcâitoare care îţi amintesc de responsabilităţile pe care le ai.

Alte programe întâmpină aceleaşi probleme. Cu toate că maşinile Über fără şofer au nevoie de intervenţie umană la fiecare 20 de kilometri56, a-i face pe şoferi să rămână atenţi la drum este un adevărat calvar. Pe 18 martie 2018, un vehicul Über autonom a lovit mortal un pieton. Înregistrarea video realizată din interiorul maşinii a arătat că „monitorul uman” care stătea în spatele volanului nu se uita la drum în minutele imediat premergătoare coliziunii.57

Aceasta este o problemă serioasă, însă există şi o soluţie alternativă. Companiile producătoare de maşini ar putea să accepte ideea că oamenii sunt oameni şi să recunoască faptul că atenţia ne poate fi distrasă. În definitiv, să poţi să citeşti o carte în timp ce conduci face

parte din ceea ce ne atrage, ne *seduce*, la maşinile care se conduc singure. Este diferenţa principală dintre nivelul 2 (*hands off*) şi nivelul 3 (*eyes oft*).

Cel din urmă presupune o provocare tehnică superioară faţă de nivelul 2, însă unii producători auto au început deja să-şi construiască maşinile ţinând cont de lipsa noastră de atenţie. Tehnologia Audi *traffic-jam pilot* este un asemenea exemplu.58 Aceasta poate să preia complet controlul autovehiculului atunci când rulezi în condiţii de trafic cu viteză redusă pe şosea, permiţându-ţi să te relaxezi şi să te bucuri de călătorie. Trebuie doar să fii pregătit să intervii dacă apare ceva neprevăzut.[61]

Există un motiv pentru care Audi şi-a limitat sistemul la traficul cu viteză redusă de pe drumurile cu acces limitat. Riscul unor catastrofe este mai mic într-un trafic aglomerat. Iar acesta este un lucru important. Pentru că, în momentul în care omul încetează să mai supravegheze drumul, se poate genera cea mai proastă combinaţie de împrejurări posibilă atunci când apare o situaţie de urgenţă.

Un şofer care nu este atent va dispune de foarte puţin timp ca să evalueze situaţia din jur şi să decidă ce are de făcut. Imaginaţi-vă că vă aflaţi într-o maşină care se conduce singură, auziţi un semnal de alarmă şi ridicaţi ochii din carte tocmai la timp ca să vedeţi în faţă un camion care-şi pierde încărcătura pe drumul pe care sunteţi. Într-o fracţiune de secundă trebuie să procesaţi informaţia din jur: motoreta de pe banda stângă, furgoneta din faţă care frânează brusc, maşina din unghiul mort din partea dreaptă. Vă veţi trezi că nu sunteţi deloc familiarizat cu drumul chiar în momentul în care ar trebui să-l cunoaşteţi cel mai bine; adăugaţi la asta lipsa de experienţă practică şi veţi fi cât se poate de slab pregătit

61 Fiind cu un pas înainte faţă de automatizarea parţială, vehiculele de nivel 3, cum este Audi dotat cu traffic-jam pilot, pot să preia controlul în unele situaţii, dacă sunt îndeplinite anumite condiţii. Şoferul încă mai trebuie să fie pregătit să intervină atunci când automobilul se confruntă cu un scenariu pe care nu îl înţelege,

pentru a face faţă situaţiilor care cer cel mai ridicat nivel de îndemânare.

Acest fapt a fost confirmat şi în simulările cu maşini fără şofer. Un studiu în care li se permitea participanţilor să citească o carte sau să se joace pe telefonul mobil în timp ce maşina se conducea singură a descoperit că a fost nevoie de până la 40 de secunde după declanşarea unei alarme pentru ca oamenii să recâştige controlul corespunzător asupra vehiculului.59 Este exact ce s-a întâmplat şi cu zborul 447 al Air France. Căpitanului Dubois, care ar fi trebuit să poată să salveze cu uşurinţă avionul de la prăbuşire, a avut nevoie de aproape un minut ca să-şi dea seama ce se întâmpla şi să găsească soluţia simplă care ar fi rezolvat problema.60

În mod ironic, cu cât tehnologia de conducere fără şofer se îmbunătăţeşte, cu atât mai grave devin problemele care pot să apară. Un sistem de conducere pe pilot automat neglijent conceput, care declanşează alarma la fiecare 15 minute, va menţine şoferul permanent vigilent şi îl va obliga să practice şofatul în mod constant. Cele la care trebuie să fii într-adevăr atent sunt sistemele automate performante şi sofisticate, care sunt *aproape* întotdeauna demne de încredere.

De aceea Gill Pratt, care conduce institutul de cercetare al Toyota, a spus următoarele:

Cel mai rău scenariu este o maşină care să aibă nevoie de intervenţia şoferului numai o dată la 300.000 de kilometri…

Un om obişnuit, care are o maşină [nouă] la fiecare 160.000 de kilometri, nu va remarca niciodată [când sistemul automat îi va ceda controlul şoferului]. Însă, din când în când, poate o dată la fiecare două maşini pe care le voi deţine, va exista acea singură ocazie când se va auzi brusc: „bip, bip, bip, acum este rândul tău!” Iar persoana în cauză, care nu s-a mai întâlnit cu această situaţie, probabil, de ani de zile […], nu va fi deloc pregătită să-i facă faţă.61

MARILE SPERANŢE

În ciuda tuturor acestor probleme, există motive întemeiate pentru a ne dori să progresăm către un viitor cu vehicule care se conduc singure. Binele continuă să învingă răul. Conducerea maşinii rămâne una dintre cele mai importante cauze ale deceselor care ar putea fi prevenite din întreaga lume. Dacă tehnologia este capabilă, oricât de puţin, să reducă numărul deceselor pe şosele, ai putea susţine că ar fi lipsit de etică să *nu* o pui în practică.

Şi mai există avantaje, şi nu puţine: chiar şi sistemele simple de asistare la conducere pot contribui la reducerea consumului de combustibil62 şi la descongestionarea traficului.63 în plus – haideţi să fim sinceri! – ideea de a-ţi lua mâinile de pe volan fie şi numai pentru o clipă, atunci când conduci cu 100 km/h, e de-a dreptul… cool!

Revenind, însă, la avertismentele lui Bainbridge, acestea fac aluzie la o problemă care are de-a face cu modul în care este concepută tehnologia de conducere autonomă a vehiculelor în prezent

Să luăm Tesla, de exemplu, unul dintre primii fabricanţi de maşini care au scos pe piaţă un model autonom. Nu există nicio îndoială că sistemul lor a avut un impact pozitiv, făcând conducerea maşinii mai sigură pentru cei care o folosesc – nu trebuie să cauţi prea mult ca să găseşti on-line materiale video cu sistemul „Forward Collision Warning”, care recunoaşte riscul producerii unui accident chiar înaintea şoferului, declanşând o alarmă şi salvând maşina de la coliziune.64

Numai că apare o uşoară neconcordanţă între ceea ce pot să facă maşinile – în esenţă, cu nişte senzori pretenţioşi pentru parcarea laterală şi un sistem inteligent de *cruise control* – şi limbajul utilizat pentru descrierea acestora, în octombrie 2016, de exemplu, compania a anunţat că „toate automobilele Tesla produse în prezent

sunt dotate cu hardware *full self-driving*[6263]. După un articol din *The Verge*, Elon Musk, arhitectul companiei Tesla, a adăugat: „De acum înainte, actualizarea autonomiei complete va constitui o procedură standard la toate autovehiculele Tesla”. 65 Iar conceptul de „autonomie completă” vine, fără îndoială, în conflict cu avertizarea pe care conducătorii auto trebuie să o accepte înainte de a folosi funcţia autopilot actuală: „Trebuie să menţineţi controlul asupra vehiculului dumneavoastră şi să fiţi responsabil de acţiunile acestuia”. 66

Speranţele îşi au importanţa lor. Este posibil să nu fiţi de acord, însă eu cred că o consecinţă inevitabilă a folosirii unui limbaj ce induce în eroare de către reprezentanţii unui brand de încredere este tocmai apariţia acelor persoane care îşi îndeasă portocale în volan – sau, mai rău, aşa cum am descoperit prin cotloanele mai întunecate ale web-ului, care concep şi vând dispozitive care „[permit] adoptanţilor timpurir să [conducă în timp cel opresc avertizările de verificare ale autopilotului sau le reduc volumul”.

Desigur că din toată industria de automobile, Tesla nu este singura vinovată. Orice companie din lumea asta se foloseşte de fanteziile noastre ca să-şi vândă produsele. În ceea ce mă priveşte, însă, există o diferenţă între a cumpăra un parfum pentru că sunt de părere că mă va face mai atrăgătoare şi a cumpăra o maşină deoarece consider că sistemul ei complet autonom mă va menţine în siguranţă.

Lăsând la o parte strategiile de marketing, nu pot

---

62 în februarie 2018, când scriu această carte, hardware-ul full self-

63d riving este un pachet opţional care poate fi achitat la cumpărare, cu toate că vehiculul încă nu deţine software-ul care să permită deplasări cu o conducere complet automată. Pe site-ul Tesla se precizează: „Nu putem şti cu precizie când va deveni disponibil fiecare element funcţional descris mai sus.” Vezi https://www.tesla.com/en_GB/blog/ all-tesla-cars-being-produced-now-have-full-self-driving-hardware (link accesat cel mai recent la 18 martie 2019) (n. a.).

decât să mă întreb dacă nu cumva am avut tot timpul o impresie greşită despre vehiculele fără şofer.

Ştim de acum că oamenii se pricep să înţeleagă subtilităţile, să analizeze contextul, să se bazeze pe experienţa acumulată

— Un adoptant timpuriu sau adoptant precoce este unul dintre primii clienţi ai unei companii, unul dintre primii consumatori ai unui produs sau unul dintre primii utilizatori ai unei tehnologii (cf. articolul din Wikipedia) (n.tr.).

** Este un produs care există în realitate, se numeşte *The Autopilot Buddy* şi îl poţi cumpăra la preţul „de nimic" de 179 de dolari. Este demn de menţionat că pe website se precizează, cu litere mici: *Autopilot Buddy* nu trebuie utilizat niciodată pe străzile publice". Vezi http's: //www.autopilotbuddy.com/ (link accesat cel mai recent la 19 martie 2019) (n.a.).

şi să observe tipare, patternuri. Nu ne pricepem mai deloc să fim atenţi, precişi, consecvenţi şi pe deplin conştienţi de ceea ce ne înconjoară. Deţinem, pe scurt, abilităţi şi capacităţi 100% opuse faţă de acelea ale algoritmilor.

De ce nu am urma, aşadar, exemplul software-ului din medicină de depistare a tumorilor, lăsând abilităţile maşinii să vină în completarea celor umane şi făcând, astfel, ca abilităţilor amândurora să se dezvolte? Până când vom ajunge la autonomia completă, de ce nu am răsturna ecuaţia cu susul în jos şi nu am tinde spre un sistem autonom care să vină în sprijinul conducătorului auto, iar nu viceversa? De ce nu am folosi o plasă de siguranţă, cum este ABS-ul sau sistemul de control al tracţiunii, care monitorizează cu răbdare drumul şi rămâne în alertă, pândind vreun pericol pe care şoferul l-a trecut cu vederea. Prin urmare, să folosim sistemul computerizat nu ca pe un conducător auto, ci ca pe un supraveghetor.

Aceasta este ideea care stă la baza muncii depuse de institutul de cercetări al Toyota. Cei de acolo includ în

construcţia maşinii două sisteme. Sistemul *chauffeur*[64], care – asemănător tehnologiei *troffic-jam pilot* de la Audi – ar putea prelua controlul într-un trafic foarte aglomerat, şi sistemul *guardian*[65], care rulează în fundal atunci când conduce şoferul uman şi acţionează ca o plasă de protecţie67, reducând riscul unui accident dacă se iveşte ceva ce şoferul nu a observat.

Compania Volvo a adoptat o strategie similară. Sistemului său de frânare automată în caz de urgenţă[66], care încetineşte automat automobilul dacă se apropie prea mult de acela din faţa sa, i se recunosc pe scară largă meritele pentru impresionantul record de siguranţă al lui Volvo XC90. Din anul 2002, când maşina a apărut pentru prima oară pe piaţa din Marea Britanie, au fost achiziţionate peste 50.000 de vehicule şi niciun şofer sau pasager din interiorul lor nu a fost ucis în vreo coliziune.68

Ca şi în cazul tehnologiei „iară şofer", despre care se discută cu atâta înflăcărare, va trebui să aşteptăm şi să vedem cum se vor termina toate acestea. Însă un lucru este sigur: cu timpul, conducerea autonomă ne va învăţa câteva lecţii care vor putea fi aplicate şi în afara universului automobilistic. Nu numai despre cum să transferăm controlul, ci şi cum să dăm dovadă de realism când ne formulăm aşteptările despre ceea ce pot algoritmii să facă.

Pentru ca acest demers să fie viabil, trebuie să ne schimbăm felul de a gândi şi să renunţăm la ideea că automobilele trebuie să funcţioneze perfect de fiecare dată şi să acceptăm că, deşi defecţiunile mecanice ale vehiculelor vor deveni un eveniment rar întâlnit, mai avem drum lung de străbătut până când „defectarea" algoritmilor să fie şi ea la fel de rară.

Aşadar, ştiind că greşelile sunt inevitabile, ştiind că,

---

64 „Şofer" (fr. în original) (n. tr.).

65 „Păzitor, supraveghetor" (engl.) (n. tr.).

66 Autonomous Emergency Breaking (AEB) (engl.), „sistem de frânare autonom" (n. tr.).

în cazul în care vom merge mai departe, nu avem altă alternativă decât să acceptăm incertitudinea, dilemele morale din lumea maşinilor fără şofer ne vor determina să decidem cât de bună trebuie să fie o astfel de maşină înainte de a fi dispuşi să o lăsăm să circule liberă pe străzile noastre. Este o întrebare importantă, iar ea rămâne valabilă şi într-un context diferit. Cât de bun înseamnă „destul de bun”? Odată ce ai construit un algoritm defectuos care *poate* să calculeze ceva, ar trebui să-l şi laşi să o facă?

Capitolul 6

9

INFRACŢIONALITATEA

Era o zi călduroasă de iulie din 1995, când o studentă de 22 de ani şi-a strâns cărţile, a părăsit biblioteca din Leeds şi s-a îndreptat spre maşina ei. Îşi petrecuse ziua punând la punct ultimele detalii ale lucrării de diplomă, iar acum era liberă să se bucure de restul verii. Tocmai când se aşeza în maşină la volan, a auzit pe cineva alergând în direcţia ei prin parcarea supraetajată. Înainte de a putea să reacţioneze, un bărbat s-a aplecat spre ea prin geamul deschis al maşinii şi i-a pus cuţitul la gât. A silit-o să se aşeze pe bancheta din spate, a legat-o, i-a lipit pleoapele cu lipici, a trecut la volan şi a plecat cu maşina.

După o călătorie îngrozitoare, bărbatul a oprit maşina pe un taluz năpădit de buruieni. Fata a auzit un zgomot când necunoscutul şi-a rabatat scaunul, iar după aceea un foşnet când a început să-şi desfacă şliţul. Ştia că intenţiona să o violeze. Luptându-se orbeşte, fata şi-a ridicat genunchii la piept şi apoi i-a împins în faţă cu toată puterea, forţându-l să se dea înapoi. În timp ce ea îl lovea cu picioarele şi se zbătea, bărbatul s-a tăiat la degete cu cuţitul pe care-l ţinea în mână, iar sângele i-a picurat pe scaunele maşinii. A lovit-o pe tânără de două ori peste faţă, apoi, spre imensa ei uşurare, bărbatul a ieşit din maşină şi a plecat. La două ore de la începerea calvarului, studenta a fost găsită rătăcind pe Globe Road din Leeds, ameţită şi cu

părul despletit, cu cămaşa sfâşiată, cu faţa înroşită acolo unde bărbatul o lovise şi cu pleoapele lipite cu lipici.1

Agresiunile sexuale asupra persoanelor necunoscute, ca în acest caz, sunt extrem de rare, însă, atunci când au loc, ele par să facă parte dintr-o serie. Mai mult ca sigur, nu era prima oară când bărbatul acela comitea o asemenea faptă. Când poliţia a analizat sângele din maşină, a descoperit că ADN-ul bărbatului era identic cu acela dintr-o probă prelevată de la locul unde fusese comis un viol cu doi ani în urmă, într-o altă parcare supraetajată. Acel atac avusese loc în Nottingham, cu vreo sută de kilometri mai spre sud. Iar, ca urmare a unui apel lansat în programul *Crimewatch* al BBC, poliţia a reuşit să lege cazul de alte trei incidente, petrecute cu zece ani înainte, în Bradford, Leeds şi Leicester.2

Numai că găsirea acestui violator în serie nu avea să fie uşoară. Toate infracţiunile fuseseră comise pe o arie de 7 046 km² – o suprafaţă imensă din Marea Britanie. În plus, ele le ofereau poliţiştilor un număr uluitor de mare de potenţiali suspecţi – 33 628 în total – fiecare din ei trebuind să fie eliminat din cercetările poliţiei sau investigat.3

Se impunea, aşadar, organizarea unei operaţiuni de căutare de proporţii uriaşe, ceea ce nu se întâmpla pentru prima oară, căci, după agresiunile din urmă cu zece ani, fusese întreprinsă o razie de proporţii, numai că, în ciuda faptului că poliţiştii bătuseră la 14 153 de uşi şi recoltaseră numeroase probe biologice (fire de păr, beţişoare cu probe de salivă pentru prelevarea ADN-ului şi altele), ancheta nu dusese, în cele din urmă, nicăieri. Era un risc foarte mare ca şi această nouă căutare să aibă aceeaşi soartă, până când Kim Rossmo, un fost poliţist canadian, şi algoritmul pe care tocmai îl crease au fost chemaţi în ajutor.4

Rossmo avusese o idee îndrăzneaţă: în loc să ţină cont de imensul volum de dovezi deja existente, algoritmul lui ignora, practic, totul. În schimb, îşi concentra atenţia

exclusiv asupra unui singur factor: localizarea geografică.

Rossmo considera că, poate, un făptaş nu alege la întâmplare *locul* unde îşi atacă victima. Poate că selectarea acestuia nu este o decizie luată absolut liber sau conştient. Chiar dacă agresiunile avuseseră loc de-a lungul şi de-a latul Angliei, Rossmo s-a întrebat dacă nu cumva ar putea exista un tipar involuntar ascuns în repartizarea geografică a infracţiunilor – un tipar suficient de simplu, încât să poată fi exploatat. Era o şansă, credea el, ca locul în care aveau loc infracţiunile să îl trădeze pe cel de unde provenea făptaşul. Cazul violatorului în serie era un prilej de a-şi pune teoria la încercare.

OPERAŢIUNEA LYNX ŞI ASPERSORUL

Rossmo nu era prima persoană care sugera că infractorii alcătuiesc, fără să-şi dea seama, tipare geografice. Ideile lui au o istorie care ajunge până prin anii 1820, când André-Michel Guerry, un avocat devenit statistician care lucra la Ministerul de Justiţie al Franţei, a început să strângă dosarele despre violurile, crimele şi furturile care aveau loc în diverse regiuni ale ţării.5

Deşi colectarea acestui tip de statistici pare un lucru destul de obişnuit în zilele noastre, în perioada aceea, matematica şi statistica fuseseră aplicate doar în cadrul ştiinţelor exacte, în care ecuaţiile erau folosite pentru a descrie în mod elegant legile fizice ale universului, pentru trasarea unor lucruri precum trasarea traiectoriei unei planete pe cer sau calcularea forţelor care acţionează într-un motor cu aburi. Până atunci nimeni nu-şi dăduse osteneala să strângă informaţii despre infracţiuni. Nimeni nu avea habar ce-ar fi trebuit să calculeze, cum să calculeze sau cu ce frecvenţă ar trebui să facă respectivele calcule. De altfel – gândeau oamenii pe vremea aceea – ce rost ar fi avut? Un bărbat era puternic, independent din fire şi se deplasa de colo-colo, acţionând după propria voinţă. Purtarea sa era imposibil de surprins de neînsemnata disciplină a statisticii.6

Însă analiza făcută de Guerry adevăratului său

recensământ al infractorilor de pe teritoriul Franţei sugera cu totul altceva. Oriunde te-ai fi aflat în Franţa, a descoperit el, infracţiunile comise prezentau anumite tipare care puteau fi recunoscute: tipare despre tipul infracţiunilor, modul în care fuseseră comise şi de către cine. Persoanele tinere comiteau mai multe infracţiuni decât cele în vârstă, bărbaţii mai mult decât femeile, săracii mai mult decât bogaţii. În mod curios, a devenit în curând limpede că aceste tipare nu se modificau odată cu trecerea timpului. Fiecare regiune din Franţa avea propria statistică pe categorii de infracţiuni, care abia dacă înregistra vreo schimbare de la an la an. Cu o precizie aproape înspăimântătoare, numărul furturilor, violurilor şi omorurilor se repeta de la un an la altul. Până şi metodele folosite de ucigaşi erau predictibile. Asta însemna că Guerry şi colegii săi puteau să aleagă o zonă şi să spună în avans, cu precizie, câte omucideri cu cuţitul, sabia, prin lapidare, strangulare sau înec se puteau aştepta să aibă loc într-un anumit an.7

Aşadar, până la urmă, poate că nu era o chestiune de liber arbitru al criminalului. Crima nu se produce la întâmplare; oamenii sunt previzibili. La aproape două secole de la descoperirea ei de către Guerry, Kim Rossmo voia să exploateze exact această predictibilitate.

Munca lui Guerry se concentrase asupra tiparelor pe care le găsise la nivel naţional şi regional, dar s-a dovedit că şi la nivel individual persoanele care comiteau infracţiuni creau tipare geografice pe care te puteai baza. Asemenea nouă, infractorii au tendinţa să rămână în teritorii care le sunt familiare. Operează local. Asta înseamnă că până şi cele mai grave infracţiuni vor fi duse la îndeplinire în apropierea locului în care îşi duce viaţa făptaşul. Iar pe măsură ce te îndepărtezi de scena crimei, şansa de a găsi sălaşul acestuia se reduce treptat8, un efect cunoscut de criminologi sub denumirea de „deteriorarea relaţiilor odată cu creşterea distanţei”[67].

67 Distance decay (engl.), vezi glosarul, p. 317 (n. red.).

Pe de altă parte, este puţin probabil ca nişte criminalii în serie să vizeze victime care locuiesc *foarte* aproape de ei, pentru a evita să atragă în mod inutil atenţia poliţiei asupra lor sau să rişte să fie recunoscuţi de vecini. De aici rezultă apariţia a ceea ce se numeşte o „zonă tampon”, care include locuinţa infractorilor, o zonă în care sunt şanse foarte mici ca aceştia să comită vreo infracţiune.9

Aceste două tipare de bază – deteriorarea relaţiilor odată cu creşterea distanţei şi zona tampon –, ascunse în localizarea geografică a celor mai grave crime, au stat la baza algoritmului lui Rossmo. Începând de la locul crimei, prins cu un ac pe o hartă, Rossmo şi-a dat seama că putea să echilibreze matematic aceşti doi factori şi să schiţeze o imagine a ariei probabile în care era posibil să locuiască infractorul.

Imaginea aceasta nu este deosebit de utilă atunci când a fost comisă o singură infracţiune. În lipsa unor informaţii suficiente pentru a merge mai departe, aşa-numitul „algoritm de geo-profilare a infractorului”[68] nu-ţi va spune cu mult mai mult decât minunatul şi demodatul bun-simţ. Dar pe măsură ce se adaugă tot mai multe infracţiuni, imaginea începe să se contureze, aducând treptat la lumină o hartă a oraşului care marchează zonele în care este cel mai probabil ca făptaşul să fie prins.

Este ca şi când infractorul ar fi un aspersor care se roteşte. Aşa cum ar fi greu să prevezi exact unde va cădea următoarea picătură de apă, nu poţi să anticipezi nici unde urmează să atace acesta. Însă, după ce apa a stropit un timp, iar pe jos au căzut numeroşi stropi, este destul de uşor să-ţi dai seama, din tiparul pe care-l alcătuiesc picăturile, unde este aşezat, cu aproximaţie, aspersorul.

Exact asta s-a întâmplat şi cu algoritmul lui Rossmo în Operaţiunea Lynx – goana după violatorul în serie.

68 Profilarea infractorului (criminal profiling, engl.) reprezintă procesul realizat de specialiştii în psihologie judiciară pentru a încerca să realizeze profilul psihologic al unui răufăcător, caracteristicile personalităţii acestuia (n. red.).

Echipa ştia acum unde avuseseră loc cele cinci crime şi mai cunoştea câteva locuri unde atacatorul folosise un card de credit furat pentru a-şi cumpăra băuturi alcoolice, ţigări şi un joc video. Doar în baza acestor localizări, algoritmul a scos în evidenţă două zone principale în care considera că era posibil să locuiască făptuitorul: Miligarth şi Killingheck, ambele situate în suburbiile oraşului Leeds.10

Între timp, în centrul de coordonare, poliţia mai dispunea de o probă importantă pentru a duce mai departe acest caz: o amprentă parţială lăsată de atacator la locul comiterii unei fapte anterioare. Era un eşantion prea mic pentru ca sistemul automat de recunoaştere a amprentelor să poată să răsfoiască baza de date a infractorilor condamnaţi în căutarea unei potriviri, aşa că orice comparaţie trebuia efectuată meticulos, cu lupa, de către un expert care să examineze cu migală fiecare suspect pe rând. Operaţiunea dura, de acum, de aproape trei ani – în pofida eforturilor susţinute a 180 de ofiţeri din cinci departamente diferite – şi începea să-şi cam epuizeze energia. Fiecare indiciu ducea, de fiecare dată, pe o linie moartă.

Ofiţerii de poliţie au decis să verifice manual toate amprentele găsite în cele două locuri semnalate de algoritm. Au început cu Miligarth: o căutare printre fişele cu amprente păstrate în baza de date a poliţiei locale nu a oferit niciun rezultat. Au trecut la Killingheck şi, după 940 de ore de căutare minuţioasă printre fişele de acolo, poliţia a reuşit să descopere un nume: Clive Barwell.

Acesta era un bărbat de 42 de ani, căsătorit şi tatăl a patru copii, care fusese arestat pentru jaf armat în intervalul de timp dintre atacuri. În acel moment, lucra ca şofer de camion şi în cursul activităţii făcea în mod regulat curse lungi prin toată ţara; dar locuia în Killingheck şi îşi vizita des mama, care locuia în Miligarth, cele două zone semnalate de algoritm.11 Amprenta parţială nu fusese de ajuns, ea singură, ca să-l identifice în mod concludent, însă

un test ADN ulterior a dovedit că Barwell comisese toate acele infracţiuni odioase. Poliţiştii îl găsiseră pe infractor! Barwell şi-a recunoscut vinovăţia la proces, în octombrie 1999. Judecătorul l-a condamnat la opt condamnări pe viaţă succesive.12

Când totul s-a terminat, Rossmo a avut prilejul să facă un bilanţ al eficacităţii algoritmului. Acesta nu indicase niciodată propriu-zis numele lui Barwell, dar semnalase pe hartă zonele în care poliţia trebuia să-şi concentreze atenţia. Dacă aceasta ar fi folosit algoritmul ca să-şi organizeze lista de suspecţi în ordinea priorităţii, în funcţie de unde locuia fiecare – verificând amprentele şi luând mostre de ADN, pe rând, de la fiecare dintre ei –, nu ar mai fi trebuit să deranjeze un număr atât de mare de oameni nevinovaţi. L-ar fi găsit pe Clive Barwell după ce ar fi căutat în numai 3% din locurile din întreaga zonă.13

Algoritmul se dovedise, cu siguranţă, eficient şi se vădea a avea şi alte aspecte pozitive. Deoarece nu făcea altceva decât să organizeze, în ordinea priorităţii, o listă cu suspecţi deja existentă, nu putea fi afectat de acel tip de tendenţiozitate pe care l-am întâlnit la capitolul „Justiţia” şi nici nu avea cum să înlocuiască munca temeinică de investigare, ci doar să o eficientizeze; prin urmare, există puţine şanse ca oamenii să-i acorde vreodată o încredere exagerată.

Totodată, algoritmul acesta este şi foarte flexibil. După Operaţiunea „Lynx”, el a fost folosit de peste 350 de agenţii de combatere a criminalităţii din întreaga lume, inclusiv de FBI-ul american şi de Royal Canadian Mounted Police[69], iar informaţiile pe care le oferă depăşesc sfera infracţionalităţii (spre exemplu, a fost folosit pentru depistarea bălţilor cu apă stătută care le slujesc ţânţarilor drept locuri de reproducere, pe baza localizării cazurilor de malarie din Egipt) 14. Un doctorand de-al meu de la University College London utilizează, în prezent, acest algoritm în încercarea sa de a prezice amplasările

69 Poliţia Regală Canadiană Călare (engl.) (n. tr.).

fabricilor de bombe, pe baza locurilor în care sunt folosite dispozitive explozive improvizate. Iar nişte matematicieni din Londra l-au folosit chiar şi pentru a încerca să-l găsească pe Banksy15, artistul solitar, pornind de la locurilor în care au fost identificate graffiti de-ale sale.

Infracţiunile pentru rezolvarea cărora geo-profilarea dă cele mai bune rezultate (violuri în serie, omucideri şi atacuri violente) sunt, din fericire, rare. În realitate, marea majoritate a infracţiunilor nu presupun genul de vânătoare de oameni de care a fost nevoie în cazul lui Clive Barwell. Dacă algoritmii ar trebui să aducă o schimbare în combaterea criminalităţii şi în afara acestor situaţii extreme, atunci ar avea nevoie de un tipar geografic diferit pentru a continua să dea roade. Unul care să poată fi aplicat unui oraş ca întreg şi care să poată identifica patternurile şi ritmurile unei străzi sau ale unui colţ de stradă, pe care fiecare agent de poliţie le cunoaşte instinctiv. Din fericire, Jack Maple avea exact ceea ce trebuie.

HĂRŢI DESPRE VIITOR

În anii 1980, mulţi newyorkezi ar fi stat destul de mult pe gânduri înainte de a se încumeta să meargă cu metroul, căci acesta nu era prea frumos. Toate suprafeţele erau acoperite de graffiti, vagoanele duhneau a urină stătută, pe peroane erau o mulţime de persoane care luau droguri, iar furturile şi tâlhăria erau la ordinea zilei. Aproximativ 20 de oameni nevinovaţi erau ucişi în subteran în fiecare an, ceea ce făcea din metrou un spaţiu la fel de periculos ca oricare altul aflat la suprafaţă de oriunde din lume.

Acesta era decorul în care îşi desfăşura activitatea de ohţer de poliţie Jack Maple, care fusese promovat de puţin timp la gradul de locotenent în cadrul Tranzit Bureau[70], iar

70 Denumirea unuia dintre cele douăzeci de birouri ale NYPD (New York Police Department, Departamentul de Poliţie al New York- ului), responsabil cu siguranţa celor aproximativ 5,6 milioane de pasageri care călătoresc zilnic pe cele 25 de rute ale metroului din oraşul american (n. red.).

în anii petrecuţi în poliţie se săturase doar să reacţioneze când erau comise infracţiuni, în loc să lupte pentru combaterea acestora. Din frustrare a luat naştere o idee strălucită.

Iată ce i-a declarat Maple unui jurnalist în 1999:

Pe un perete, pe mai bine de 16 m^2, am trasat fiecare staţie de metrou din New York şi fiecare tren în parte, apoi am folosit carioci pentru a marca fiecare infracţiune violentă, tâlhărie şi furt calificat care avuseseră loc acolo. Am trecut pe hartă atât cazurile rezolvate, cât şi pe cele nerezolvate.16

Poate că nu pare cine ştie ce, însă hărţile acestea, mâzgălite pe hârtie maro cu carioca, au devenit cunoscute drept „Hărţile despre viitor" şi au fost considerate revoluţionare la timpul lor. Nimeni nu se gândise până atunci să încerce să analizeze infracţiunile în felul acela. Dar, în momentul în care Maple s-a dat puţin înapoi, să aibă o privire de ansamblu asupra infracţiunilor din întregul oraş în acelaşi timp, şi-a dat seama că privea totul dintr-o perspectivă complet diferită şi, după cum explica tot el:

Se punea întrebarea: „de ce?". Care erau cauzele responsabile de faptul că se săvârşeau o mulţime de infracţiuni într-o anumită zonă?

Problema era că, pe atunci, fiecare apel de urgenţă la poliţie era tratat ca un incident izolat. Dacă sunai să anunţi existenţa unui grup de dealeri de droguri agresivi care-şi făceau veacul la un colţ de stradă, iar aceştia se făceau nevăzuţi imediat ce veneau poliţiştii, nu se consemna nimic care să facă legătura între plângerea ta şi alte apeluri de urgenţă primite în momentul în care gaşca revenea la locul său obişnuit. Hărţile lui Maple, în schimb, îi ofereau posibilitatea să determine cu precizie locul unde infracţiunile reprezentau o problemă cronică, iar acest lucru însemna că putea să descopere şi motivele acestora:

Aici este un centru comercial? De aceea avem o mulţime de hoţi de buzunare şi de jafuri? Aici este o

şcoală? De aceea avem o problemă la ora trei după-amiaza? în vecinătate se află vreo casă abandonată? De aceea se face trafic de droguri la colţul străzii? 17

Capacitatea de a răspunde la aceste întrebări era primul pas spre soluţionarea problemelor din oraş. Aşa că, în 1990, când Bill Bratton, un tip deschis la minte, a devenit şeful NYPD, Mapple i-a arătat „hărţile despre viitor" pe care le concepuse. Le-au folosit amândoi pentru a încerca să facă metroul un loc mai sigur pentru toată lumea.18

Şi Bratton a avut, la rândul său, o idee bună. Ştia că persoanele care cerşeau sau urinau în subteran, precum şi acelea care săreau peste sistemul de tumichete reprezentau o mare problemă în staţiile de metrou. El a decis să le ceară poliţiştilor lui să îşi concentreze atenţia asupra rezolvării acestor delicte minore înainte de a soluţiona infracţiunile mult mai grave, precum jafurile şi omorurile, care erau comise, de asemenea, în proporţii catastrofale în subteran.

Raţionamentul său avea două aspecte. În primul rând, manifestând asprime faţă de orice comportament antisocial chiar la locul faptei, se transmitea în cel mai puternic mod mesajul că activitatea infracţională nu era acceptată sub nicio formă, în speranţa ca astfel să înceapă să transforme percepţia oamenilor despre ceea ce era „normal" să se petreacă în metrou. În al doilea rând, persoanele care evitau să plătească taxa de intrare era foarte posibil să fie tocmai infractorii care aveau să comită, ulterior, infracţiuni mai grave şi, dacă erau arestate pentru că nu plătiseră taxa, nu ar mai fi avut prilejul să le ducă la îndeplinire. După cum declara Bratton ziarului *Newsday* în 1991:

Prinzându-i pentru că nu plăteau taxa de acces, am reuşit să oprim la tumichete infractori periculoşi şi înarmaţi înainte ca aceştia să ajungă în metrou şi să facă ravagii.19

Strategia a funcţionat. Cu cât măsurile poliţiei

deveneau mai inteligente, cu atât metrourile erau mai sigure. Între 1990 şi 1992, hărţile lui Maple şi tacticile lui Bratton au redus fărădelegile cu 27%, iar jafurile cu o treime.20

Când Bratton a devenit comisar al Departamentului de Poliţie al oraşului New York, a decis să ia cu el „hărţile despre viitor" ale lui Maple. Acolo, acestea au fost dezvoltate şi îmbunătăţite, devenind Compătat, un instrument de urmărire a informaţiilor care este utilizat în prezent de numeroase inspectorate de poliţie, atât din Statele Unite, cât şi din străinătate. La baza sa se află tot principiul simplu al lui Jack Maple: consemnează unde au avut loc infracţiunile pentru a descoperi cele mai periculoase zone din oraş.

Aceste puncte fierbinţi tind să fie, relativ, concentrate, în Boston, de exemplu, un studiu desfăşurat pe o perioadă de 28 de ani a descoperit că 66% dintre toate jafurile stradale aveau loc pe doar 8% din străzi.21 într-un alt studiu, s-au marcat pe hartă 300.000 de locuri semnalate de apelurile de urgenţă la poliţia din Minneapolis: jumătate dintre ele proveneau de pe o suprafaţă de numai 3% din întreaga întindere a oraşului.22

Dar aceste puncte fierbinţi nu rămân neschimbate de-a lungul timpului. Ele se deplasează în mod constant, variind şi transformându-se subtil, asemenea picăturilor de ulei în apă, chiar şi de la o zi la alta. Iar când Bratton s-a mutat la Los Angeles, în 2002, a început să se întrebe dacă mai existau şi alte tipare, care ar fi putut să te anunţe *când* urma să aibă loc o infracţiune, nu numai *unde* avea să fie comisă aceasta. Exista vreo modalitate de a interpreta altfel infracţiunile care fuseseră deja comise? în loc să reacţioneze doar după săvârşirea fărădelegilor – motivul frustrării lui Maple – sau să le combată, oare nu exista vreun mod de a le prevedea?

SEMNALUL ŞI INTENSIFICAREA

Când se pune problema de a prezice infracţiunile, cel mai bine este să se înceapă de la analizarea furturilor prin

efracţie. Din moment ce acestea au loc la o anumită adresă, se ştie cu precizie unde s-au petrecut – spre deosebire de furtul din buzunare, spre exemplu (pentru că, în fond, este posibil ca victima să nu observe lipsa telefonului mobil decât atunci când ajunge acasă). Cei mai mulţi oameni vor face plângere la poliţie dacă au fost tâlhăriţi, aşa că dispunem de seturi de date foarte detaliate şi de bună calitate, care sunt mult mai greu de cules în cazul infracţiunilor legate de traficul de droguri, de exemplu. În plus, oamenii ştiu adesea destul de bine când le-au fost sparte locuinţele (poate atunci când erau la serviciu sau când şi-au petrecut seara în oraş), informaţii de care nimeni nu va dispune în cazul unor infracţiuni precum faptele de vandalism.

Spărgătorii de locuinţe au, de asemenea, ceva în comun cu criminalii şi violatorii în serie pe care i-a studiat Rossmo: ei tind „să opereze” doar în zone care le sunt familiare. Acum ştim că este mai probabil ca o locuinţă să fie spartă dacă se află pe o stradă des frecventată de un hoţ – de pildă, în drumul lui spre locul de muncă sau spre şcoală.23 Mai ştim[71] că există un punct sensibil referitor la nivelul de aglomeraţie pe care îl preferă spărgătorii: ei tind să evite străzile cu foarte mult trafic şi să le prefere pe acelea pe care se plimbă liniştiţi o mulţime de oameni care nu locuiesc în zonă, pentru că aceştia le servesc drept acoperire (atâta vreme cât nu există şi mulţi localnici băgăcioşi care-şi pierd vremea pe acolo, făcând pe paznicii improvizaţi).24

Dar acesta este numai primul din cele două elemente care fac o locuinţă atrăgătoare pentru un spărgător. Este adevărat, există şi factori care nu se schimbă în timp, cum ar fi adresa la care locuiţi sau cât de aglomerată este strada dumneavoastră, care vor „semnala”, vor atrage în mod constant atenţia asupra proprietăţii dumneavoastră,

71 De această dată, atunci când mă exprim la plural, mă refer la faptul că studiul în cauză e o lucrare pe care am realizat-o împreună cu minunatul meu doctorand, Michael Frith (n. a.).

transformând-o într-o ţintă. Dar, înainte de a vă grăbi să vă vindeţi casa şi să vă mutaţi pe o alee închisă, liniştită, şi intens patrulată de poliţia comunitară, ar trebui să mai ştiţi că în lumea infracţionalităţii punctele fierbinţi nu rămân imobile. Cel de-al doilea motiv pentru care locuinţa dumneavoastră este atrăgătoare rămâne, fără îndoială, şi cel mai important. Acest factor depinde de ceea ce se întâmplă chiar în acest moment în imediata ei vecinătate şi este cunoscut drept „intensificare”.

Dacă ţi-a fost vreodată spartă locuinţa de două ori într-un interval scurt de timp înseamnă că eşti deja bine familiarizat cu efectul de intensificare. Aşa cum îţi vor spune şi poliţiştii după ce ai devenit victima unei spargeri pentru prima oară, infractorii au tendinţa să vizeze în mod repetat acelaşi loc – ceea ce înseamnă că, oriunde ai locui, în zilele următoare jafului, există riscul de a fi prădat din nou. De fapt, şansele de a fi vizat din nou cresc de douăsprezece ori în această perioadă.25

Sunt câteva motive clare pentru care spărgătorii de locuinţe ar putea decide să revină în casa voastră. Poate au ajuns să-i cunoască planul, să afle unde vă păstraţi bunurile de valoare (obiecte precum televizoarele şi calculatoarele sunt înlocuite destul de repede), să se familiarizeze cu încuietorile uşilor sau căile de scăpare din zonă, ori poate că au observat un obiect valoros pe care nu au putut să-l ia de prima dată. Oricare ar fi motivul, acest efect de intensificare nu este valabil numai pentru dumneavoastră. Cercetătorii au descoperit că şansele ca vecinii să fie jefuiţi imediat după o spargere vor creşte de asemenea, la fel ca şi cele ale vecinilor vecinilor voştri, ale vecinilor vecinilor vecinilor voştri şi aşa mai departe, pe toată strada.

Vă puteţi imagina cum aceste intensificări îşi fac apariţia şi se aprind în tot oraşul, ca un foc de artificii. Pe măsură ce te îndepărtezi de scânteia iniţială, impulsul devine din ce în ce mai slab, iar, cu timpul, se atenuează treptat până când, după două luni – dacă nu cumva o nouă

spargere nu relansează mecanismul pe aceeaşi stradă –, va fi dispărut în totalitate.26

Semnalul şi intensificarea întâlnite în fenomenul infracţional au, de fapt, un corespondent într-un fenomen natural: cutremurele de pământ. Este adevărat, nu poţi preciza exact unde şi când va lovi primul cutremur (deşi ştii că unele regiuni sunt mai predispuse decât altele). Dar, în momentul în care se manifestă primele trepidaţii ale solului, poţi să prezici cu destul de multă siguranţă locul şi la frecvenţa replicilor care e de aşteptat să urmeze, riscul fiind cel mai mare acolo unde s-a petrecut cutremurul iniţial, scăzând pe măsură ce te îndepărtezi de el şi estompându-se odată cu trecerea timpului.

Sub conducerea lui Bratton s-a făcut pentru prima oară legătura între structura unui cutremur şi spargerea locuinţelor. Dornic să găsească o cale de a prevedea comiterea infracţiunilor, Departamentul de Poliţie din Los Angeles a încheiat un parteneriat cu un grup de matematicieni de la UCLA[72] şi i-a lăsat pe aceştia să-şi bage nasul prin toate informaţiile pe care poliţiştii reuşiseră să pună mâna: 13 milioane de infracţiuni comise într-o perioadă de peste 80 de ani. Deşi criminaliştii erau deja de câţiva ani la curent cu fenomenul de semnalare şi intensificare, în încercarea de a descoperi anumite patternuri în acel noian de informaţii, cercetătorii UCLA au fost primii care şi-au dat seama că ecuaţiile matematice care preziceau atât de bine riscul seismic şi replicile care apăreau în urma unui cutremur puteau fi, la fel de bine, folosite pentru a prevesti infracţiunile şi „postinfracţiunile”. Şi, în plus, că ele nu dădeau roade numai pentru spargerile din locuinţe: ei descoperiseră, practic, o modalitate de a prevedea orice, de la furtul de maşini şi până la violenţe şi fapte de vandalism.

Implicaţiile erau cu adevărat… cutremurătoare. În loc să afirmi doar că o zonă din oraş care căzuse recent

72 University of California, Los Angeles (engl., Universitatea din California, statul Los Angeles) (n. tr.).

victimă unor infracţiuni era relativ „mai expusă riscului”, cu ajutorul acestor ecuaţii puteai să cuantifici cu exactitate care era riscul respectiv, ajungând chiar până la nivelul unei singure străzi. Iar ştiind, probabilistic vorbind, că o anumită zonă din oraş urma să fie ţinta spărgătorilor de case într-o anumită noapte, matematicienilor le-a fost uşor să scrie un algoritm care să avertizeze poliţia unde să-şi concentreze atenţia.

Şi aşa a luat naştere Pred Pol (sau PRED – ictive POL – icing[73]).

INFORMATORUL

Se prea poate să fi auzit deja de Pred Pol. De la lansare, în 2011, a făcut subiectul a mii de articole de ziar, de obicei, sub un titlu care făcea referire la filmul *Minority Report* cu Tom Cruise. A devenit un fel de Kim Kardashian al algoritmilor: foarte celebru, extrem de criticat în mass-media, deşi nimeni nu pare să înţeleagă cu adevărat ce face.

Aşa că, înainte să vă umpleţi mintea cu imaginea unor clarvăzători care stau într-un bazin cu apă şi îşi strigă cu voce tare premoniţiile, daţi-mi voie să vă înşel puţin aşteptările; Pred Pol nu urmăreşte oamenii înainte ca aceştia să comită o infracţiune; nu poate, în niciun caz, să vizeze un anumit individ, ci doar zone geografice. Ştiu că am utilizat în mod repetat cuvântul „predicţie”, însă algoritmul nu poate, de fapt, să prezică exact ce se va întâmpla în viitor. Nu este un glob de cristal. Nu poate prevedea decât riscul apariţiei unor evenimente viitoare, nu şi evenimentele în sine, iar aceasta este o diferenţă subtilă, dar importantă.

Gândiţi-vă la algoritm ca la un fel de agent de pariuri. Dacă un grup numeros de ofiţeri de poliţie se înghesuie în jurul hărţii oraşului, făcând pariuri în privinţa locului în care va fi săvârşită o infracţiune în acea noapte, atunci Pred Pol calculează şansele ca acest lucru să se întâmple.

---

73 „Măsuri poliţieneşti preventive” (sau „criminalistică algoritmică”) (n. tr.).

Acţionează ca un fel de informator, scoţând în evidenţă străzile şi zonele care sunt „favoritele” din noaptea aceea, sub forma unor pătrate mici şi roşii pe o hartă.

Întrebarea-cheie este dacă merită să ţii cont de „favoritul” pe care ţi-l dezvăluie „informatorul”. Pentru a verifica dibăcia algoritmului27, acesta a fost pus să se confrunte cu cei mai buni analişti umani, specialişti în domeniul infracţionalităţii, în cadrul a două experimente diferite: unul realizat în Kent, în partea de sud a Angliei, iar celălalt, în regiunea din sud-vestul oraşului Los Angeles. Testul a constat într-o confruntare cinstită, faţă în faţă. Tot ce aveau de făcut algoritmul şi specialiştii în infracţionalitate era să plaseze pe o hartă 20 de pătrate, fiecare reprezentând o suprafaţă de 150 m^2, pentru a indica unde credeau ei că avea să aibă loc un număr mai mare de infracţiuni în următoarele 12 ore.

Înainte de a vedea care au fost rezultatele, este important să subliniem cât de dificil este acest lucru. Dacă mie sau dumneavoastră ni s-ar da aceeaşi însărcinare, presupunând că nu avem o cunoaştere aprofundată a zonelor cu o rată mare de infracţiuni din Kent sau California, probabil că nu am face decât să aşezăm la întâmplare pătratele pe hartă. Acestea ar acoperi o porţiune infimă din ea – numai o miime din suprafaţa totală a comitatului Kent28 –, iar la fiecare 12 ore ar trebui să eliminăm toate prezumţiile anterioare şi să o luăm de la capăt, împrăştiind pătratele la întâmplare, ne-am putea aştepta „să prezicem” cu succes mai puţin de 1% din infracţiuni.29

Specialiştii s-au descurcat mult mai bine. În Los Angeles, analistul a reuşit să prezică în mod corect unde vor avea loc 2, 1% din infracţiuni30, iar specialiştii britanici au obţinut un procent şi mai bun, cu o medie de 5, 4%31, un scor absolut impresionant, ţinând cont de faptul că harta lor era cam de zece ori mai mare decât aceea a oraşului Los Angeles.

Însă algoritmul i-a eclipsat pe toţi. În Los Angeles a

prezis corect mai mult decât dublul infracţiunilor pe care reuşiseră să le prezică specialiştii umani, iar la o etapă a testului din Marea Britanie, aproape una din cinci infracţiuni s-a produs în pătratele roşii pe care le prevăzuseră ecuaţiile matematice.32 Pred Pol nu este un glob de cristal pentru ghicit, dar niciuna dintre metodele folosite în trecut nu a reuşit să prevadă infracţionalitatea viitoare cu un succes atât de mare.

UTILIZAREA PREDICŢIILOR

Numai că apare o problemă. Deşi algoritmul este destul de bun la *prezicerea* teritoriului în care vor avea loc infracţiuni în următoarele 12 ore, poliţia are un obiectiv puţin diferit: *reducerea* infracţionalităţii în următoarele 12 ore. Odată ce algoritmul a furnizat predicţiile, nu este foarte clar ce ar trebui făcut în continuare.

Desigur, există câteva opţiuni. În cazul spargerilor de locuinţe, s-ar putea instala camere CCTV sau s-ar putea aduce poliţişti sub acoperire care să-i prindă pe răufăcători în flagrant delict. Sau ar fi mai bine pentru toată lumea dacă toate eforturile s-ar îndrepta spre prevenirea infracţiunilor înainte ca acestea să se întâmple. În definitiv, ce aţi prefera?

Să fiţi victima unei infracţiuni al cărei făptaş va fi prins sau să nu fiţi deloc victima unei infracţiuni?

De asemenea, ar putea fi avertizaţi locuitorii din zonă că bunurile le sunt în pericol, poate chiar să li se ofere sprijin de către poliţie în îmbunătăţirea încuietorilor sau s-ar putea instala sisteme de alarmă sau temporizatoare la întrerupătoarele luminilor pentru a păcăli indivizii îndoielnici care ar trece pe stradă, făcându-i să creadă că în locuinţe este mereu cineva. Este ceea ce s-a întreprins în cadrul unui experiment din 2012 în Manchester33, unde s-a reuşit reducerea numărului de furturi prin efracţie cu peste o treime. A existat un mic dezavantaj, însă: cercetătorii au calculat că această tactică, numită şi „asigurarea siguranţei ţintei”, a costat aproximativ 3 925 lire pentru fiecare spargere de locuinţă preîntâmpinată.34

Şi nu v-aş sfătui să încercaţi să convingeţi Departamentul de Poliţie din Los Angeles să pună în aplicare ideea aceasta, căci ofiţerii săi se confruntă cu peste 15.000 de spargeri de locuinţe pe an! 35

O altă opţiune, care se abate cât mai puţin posibil de la metodele poliţieneşti tradiţionale, este tactica numită *cops on the dots*. Iată ce mi-a declarat Steyve Coloan, un ofiţer de poliţie englez ieşit la pensie care a lucrat la Metropolitan Police**:

În zilele de demult – patrulele erau repartizate doar pe zone geografice. Luai harta, o fragmentai pe zone şi pe acestea le împărţeai pe sectoare. „Tu mergi în sectorul ăla, tu mergi în sectorul ăla!” Simplu ca bună ziua.

Problema era, aşa cum a reieşit dintr-un studiu realizat în Marea Britanie, că un ofiţer de poliţie care patrula pe jos prin sectorul care îi fusese alocat la întâmplare se putea aştepta să

* „Poliţişti [picaţi] la ţanc” (într-o traducere aproximativă) (n.tr.). ** „Poliţia Metropolitană” (engl.), denumire sub care este cunoscută poliţia londoneză (oficial: Metropolitan Police Force sau Met), care acţionează în cele 32 de burguri londoneze, în afara City-ului, unde operează City of London Police (n. red.).

ajungă la o sută de metri de locul unde avea loc o spargere de locuinţă numai o singură dată în 80 de ani.36

Folosind metoda *cops on the dots*, în schimb, patrulele erau trimise în punctele fierbinţi pe care le semnala algoritmul. (Ceea ce înseamnă că metoda ar fi trebuit să se numească, atunci, *cops on the hotspots*[74], numai că, probabil, denumirea nu ar fi fost la fel de captivantă.) Ideea era, desigur, că, dacă prezenţa poliţiştilor este mai vizibilă, la locul potrivit şi la timpul potrivit, e mult mai probabil să oprească desfăşurarea unei infracţiuni sau cel puţin să le permită ofiţerilor să reacţioneze rapid în perioada imediat următoare.

Este exact ce s-a întâmplat în comitatul Kent. În cea

74 „Poliţişti în punctele fierbinţi” (engl.) (n. tr.).

de-a doua etapă a studiului, când tocmai începea schimbul de noapte, sergentul a scos la imprimantă hărţile cu pătrate roşii care marcau zonele cu risc ridicat din acea seară. Când patrula de poliţie avea o perioadă mai liniştită, se ducea la cel mai apropiat pătrat roşu, cobora din maşină şi străbătea zona pe jos.

Într-o seară, într-o zonă în care, în mod normal, nu s-ar fi dus niciodată, poliţiştii au găsit pe stradă o femeie din Europa de Est şi copilul ei. A reieşit că aceasta era implicată într-o relaţie abuzivă şi că, în urmă cu doar câteva minute, copilul fusese agresat sexual. Sergentul care era de gardă în acea noapte a confirmat că „i-am găsit pe oamenii aceştia pentru că se aflau într-un pătrat Pred Por.37 Suspectul a fost arestat în apropiere de locul acela, mai târziu, în aceeaşi noapte.

Mama şi copilul ei nu au fost singurele persoane ajutate de algoritm pe parcursul experimentului *cops on the dots*: infracţionalitatea din Kent a scăzut în medie cu 4%. Studii similare întreprinse în Statele Unite (efectuate chiar de Pred Pol) au anunţat scăderi şi mai importante ale infracţionalităţii. În zona Foothill din Los Angeles infracţionalitatea a scăzut cu

13% în primele patru luni de utilizare a algoritmului, în pofida unei creşteri de 0,4% în restul oraşului, unde poliţia întrebuinţa metodele tradiţionale. Iar Alhambra, un oraş din California situat nu foarte departe de Los Angeles, a anunţat o uimitoare scădere de 32% a spargerilor de locuinţe şi de 20% a furturilor de maşini după introducerea algoritmului, în ianuarie 2013.38

Aceste cifre sunt impresionante, dar, în realitate, este greu de afirmat cu certitudine dacă meritul este al Pred Pol. Toby Davies, matematician şi criminolog de la University College London, mi-a spus:

Este posibil ca până şi simplul fapt că ofiţerii de poliţie sunt încurajaţi să se ducă în diverse locuri, să coboare din maşină şi să patruleze pe jos, indiferent unde, să fi dus oricum la reducere a criminalităţii.

Iar aici mai apare o problemă. În situaţia în care cu cât cauţi mai temeinic infracţiunile, cu atât ai mai multe şanse să le găseşti, atunci faptul de a trimite poliţia pe străzi ar putea să schimbe evidenţa infracţiunilor, căci tot Davies mi-a declarat:

Când poliţiştii sunt undeva, descoperă mai multe infracţiuni comise acolo decât ar fi făcut-o altminteri. Chiar şi atunci când se petrec, în mod egal, infracţiuni în două locuri diferite, ofiţerii vor descoperi ceva mai multe acolo unde au fost prezenţi, decât vor descoperi în viitor în acela în care nu au fost.

Asta înseamnă că probabil există un dezavantaj uriaş aş metodei *cops on the dots*: trimiţând poliţiştii într-o anumită zonă pentru a combate infracţionalitatea pe baza predicţiilor algoritmului, se riscă intrarea într-un cerc vicios. Pentru că, dacă, spre exemplu, un cartier mai modest are un nivel mai mare de criminalitate într-o primă etapă, algoritmul poate prezice că în viitor acolo se vor produce încă şi mai multe infracţiuni; prin urmare, în acel cartier vor fi trimişi ofiţerii de poliţie, iar aceştia vor descoperi mai multe infracţiuni. Implicit, algoritmul va prezice comiterea şi mai multor infracţiuni, vor fi trimişi acolo încă şi mai mulţi poliţişti şi aşa mai departe.

Acest cerc vicios care se tot repetă s-ar putea să devină o problemă pentru stoparea infracţiunilor corelate cu zonele mai sărace, cum ar fi cerşitul, vagabondajul şi consumul de droguri uşoare.

În Marea Britanie, unde unele categorii sociale se plâng în mod constant de lipsa prezenţei poliţiştilor pe străzi, concentrarea atenţiei poliţiştilor asupra anumitor zone s-ar putea să nu pară inechitabilă imediat. Însă nu toată lumea este în bune relaţii cu poliţia, căci, după cum afirma Davies:

Este firesc ca persoanele care observă că un ofiţer de poliţie se plimbă prin faţa locuinţei lor în fiecare zi să se simtă persecutate din acest motiv, chiar dacă nimeni nu comite nicio infracţiune şi chiar dacă poliţistul respectiv

nu face nimic altceva decât să parcurgă strada de la un capăt la altul. Este aproape ca şi când ai avea dreptul să nu te afli constant sub presiune, permanent supravegheat de poliţie.

Iar eu, una, tind să-i dau dreptate.

Un algoritm bine pus la punct *ar trebui* să fie astfel dezvoltat, încât să poată ţine cont de tacticile utilizate de poliţie. Teoretic, cel puţin, există modalităţi să te asiguri că algoritmul nu vizează cu precădere anumite cartiere (spre exemplu, se pot trimite, în mod aleatoriu, poliţişti în zone cu risc mediu şi ridicat de comitere a unor infracţiuni). Din păcate, însă nu există nicio modalitate de a şti cu siguranţă dacă Pred Pol reuşeşte să evite cu totul acest cerc vicios care se tot repetă sau dacă operează echitabil la nivel general, întrucât Pred Pol este un algoritm brevetat, aşadar, codul său sursă nu este accesibil publicului şi nimeni nu ştie exact cum funcţionează.

Pred Pol nu este singurul software existent pe piaţă. Un concurent al său este Hunchlab, care lucrează prin combinarea diferitelor date statistice despre o zonă: infracţiuni declarate, apeluri de urgenţă, date de recensământ (precum şi parametri mai surprinzători cum ar fi fazele lunii). Hunchlab nu se bazează pe nicio teorie. Nu încearcă să stabilească *de ce* infracţiunile au loc mai mult într-o anumită zonă decât în alta; el nu face decât să prezinte patternurile pe care le găseşte în toate acele informaţii. În consecinţă, poate prezice cu fiabilitate mai multe tipuri de infracţiuni decât Pred Pol (care are la bază teorii despre modul în care infractorii creează tipare geografice) – dar, din moment ce şi Hunchlab este protejat prin drepturi de proprietate intelectuală, este, practic, imposibil să ne asigurăm, din afară, că nu discriminează, în mod accidental, anumite segmente de populaţie.39

Un alt algoritm predictiv lipsit de transparenţă este Strategic Subiect List[75], folosit de Departamentul de Poliţie din Chicago.40 Acesta are o abordare cu totul diferită de a

75 „Lista cu subiecţi strategici” (engl.) (n. tr.).

celorlalţi: în loc să se concentreze asupra criteriilor geografice, el încearcă să prevadă ce indivizi vor fi implicaţi în infracţiuni în care se face uz de armă. Utilizând o mare diversitate de factori, algoritmul creează o „listă fierbinte" de persoane despre care consideră că e posibil să fie implicate în acte de violenţă armată în viitorul apropiat, fie împuşcând pe altcineva, fie fiind împuşcate. Teoria pe care se bazează este valabilă: victimele de astăzi sunt, adesea, făptaşii de mâine. Iar programul este bine intenţionat: ofiţerii de poliţie vizitează persoanele de pe lista de supraveghere pentru a le oferi acces la diverse programe de intervenţie şi a le ajuta să-şi schimbe viaţa.

Numai că există motive de îngrijorare că Strategic Subiect List s-ar putea să nu se ridice la înălţimea aşteptărilor. O investigaţie recentă, realizată de organizaţia non-profit RAND Corporation[76] a ajuns la concluzia că faptul de a figura pe această listă nu avea nicio importanţă din perspectiva probabilităţii ca un individ să fie implicat într-un schimb de focuri.41 însemna, totuşi, că era mai probabil să fie arestat. Poate – conchidea raportul – din cauză că ofiţerii tratau pur şi simplu lista de supraveghere ca pe o listă de suspecţi oricând avea loc un schimb de focuri.

Algoritmii predictivi utilizaţi de poliţie sunt, fără îndoială, promiţători şi oamenii responsabili de crearea lor fac, neîndoielnic, acest lucru cu toată bună credinţa şi cu cele mai bune intenţii. Însă îngrijorările privind subiectivismul şi discriminarea sunt legitime. Iar în opinia mea aceste chestiuni sunt prea importante pentru construirea unei societăţi echitabile pentru ca noi să acceptăm, pur şi simplu, asigurări precum aceea că forţele

76 Acronim pentru Research ANd Development Corporation, think-thank dedicat politicilor globale înfiinţat în 1948 de Douglas Aircraft Company (una dintre companiile aviatice care au înflorit datorită contractelor militare obţinute în perioada celui de-al Doilea Război Mondial), cu scopul declarat de a oferi capacităţi de cercetare şi analiză economică armatei SUA. în prezent, RÂND Corporation este finanţată de guvernul american şi de diverse persoane fizice şi juridice (n. red.).

de ordine le vor utiliza într-un mod nepărtinitor. Este unul dintre multele exemple care ne dau de înţeles cât de stringent avem nevoie de specialişti independenţi şi de un organism de reglementare care să se asigure că binele pe care îl fac algoritmii nu va fi depăşit de răul produs de ei.

Iar posibilele efecte dăunătoare ne depăşesc capacităţile de predicţie. Aşa cum am văzut deja într-o serie de alte exemple, există un pericol real ca algoritmii să adauge o notă de autoritate unui rezultat incorect. Iar aici consecinţele pot fi dramatice. Doar pentru că un calculator afirmă un lucru nu înseamnă că acesta este şi adevărat.

CINE TE CREZI?

În 2014, Steve Talley dormea în locuinţa lui din South Denver când a auzit o bătaie în uşă.42 A deschis, iar în pragul ei se afla un bărbat care îşi cerea scuze că îi lovise din greşeală maşina. Străinul l-a rugat pe Talley să iasă până afară, să vadă despre ce este vorba. El i-a făcut pe plac. Pe când se ghemuia, să evalueze daunele suferite de portiera din partea şoferului a maşinii sale43, s-a declanşat o grenadă cu efect luminos. Şi-au făcut apariţia trei bărbaţi îmbrăcaţi cu jachete şi căşti negre, care l-au doborât: un bărbat i-a imobilizat capul, altul – braţele, pe când un al treilea a început să-l lovească în mod repetat cu patul pistolului.

Talley s-a ales cu răni multiple. Până la sfârşitul serii, suferise leziuni ale nervilor, i se formaseră cheaguri de sânge şi se alesese cu o ruptură de penis.44 „Nici măcar nu ştiam că un penis poate fi rupt”, a declarat el mai târziu unui ziarist de la *The Intercept*. „La un moment dat strigam, pur şi simplu, să vină poliţia. Pe urmă, mi-am dat seama că aceia care mă băteau erau chiar poliţişti”. 45

Steve Talley a fost arestat pentru două jafuri comise la băncile din apropiere. În timpul celui de-al doilea, fusese agresat un ofiţer de poliţie, motiv pentru care, credea Talley, era tratat cu atâta brutalitate în timpul arestării. „Le-am spus că erau nebuni”, a declarat, amintindu-şi că

le-ar fi strigat poliţiştilor: „Aţi greşit persoana!”

Talley nu minţea. Arestarea sa era rezultatul asemănării lui uimitoare cu persoana potrivită – adevăratul jefuitor de bănci.

Un om de serviciu, angajat în clădirea în care lucra Talley, informase, iniţial, poliţia după ce văzuse fotografiile făptaşului la ştirile locale, însă un specialist FBI examinase ulterior înregistrările camerei CCTV folosind un software de recunoaştere facială46 şi ajunsese la concluzia că „individul care apare în înregistrare pare să fie Talley”.47

KI Imaginea lui Steven Talley Q1 Imaginea filmată de camera video din holul băncii americane

Talley avea un alibi de nezdruncinat, însă din cauza declaraţiei expertului FBI, avea să fie nevoie de mai bine de un an întreg înainte ca numele să-i fie pe deplin reabilitat. Între timp, a fost ţinut într-o celulă de maximă securitate aproape două luni, până s-au strâns dovezi suficiente pentru eliberarea lui. În consecinţă, a fost pus în imposibilitatea de a mai lucra, iar până când calvarul a luat sfârşit, Talley îşi pierduse deja slujba, casa şi dreptul de a-şi vedea copiii. Toate acestea ca o consecinţă directă a acestei false identificări.

ALGORITMI CARE VĂD DUBLU

Algoritmii de recunoaştere facială devin deja un instrument folosit în mod obişnuit în investigaţiile poliţieneşti moderne. Când li se prezintă o fotografie, o înregistrare sau o captură de la un aparat de fotografiat 3D, aceşti algoritmi vor detecta faţa care apare, îi vor

măsura caracteristicile şi o vor compara cu acelea din baza de date cu chipuri cunoscute pentru a determina identitatea persoanei reprezentate acolo.

La Berlin, algoritmii de recunoaştere facială capabili să identifice suspecţii de terorism cunoscuţi sunt antrenaţi cu ajutorul mulţimilor de oameni care trec prin gări.48 în SUA, numai în statul New York aceşti algoritmi au dus la peste 4.000 de arestări, începând din 2010, *doar* pentru fraudă şi furt de identitate.49 Iar în Marea Britanie, camere video instalate pe vehicule care arată ca nişte maşini care lucrează pentru Google Streetview străbat în prezent oraşele, verificând, în mod automat, dacă semănăm sau nu cu indivizi din baza lor de date cu persoane căutate.50 Aceste dube au obţinut primul succes în iunie 2017, după ce au trecut pe lângă un bărbat din South Wales, pe numele căruia fusese emis un mandat de arestare.51

Siguranţa şi securitatea depind adesea de capacitatea noastră de a identifica şi de a recunoaşte chipurile umane.

Însă a lăsa această sarcină în seama oamenilor poate fi riscant. Să-i luăm, de exemplu, pe ofiţerii de la paşapoarte. Într-un studiu recent, menit să reproducă toate condiţiile de securitate dintr-un aeroport, aceşti profesionişti ai recunoaşterii chipurilor nu au reuşit să observe persoane care aveau acte de identitate greşite în uimitorul procent de 14% din cazuri – şi au respins în mod incorect 6% dintre persoanele care aveau un act de identitate absolut corespunzător.52 Nu ştiu ce părere aveţi voi, însă eu găsesc că aceste procente sunt mai mult decât îngrijorătoare, dacă stau să mă gândesc la numărul de persoane care trec zi de zi prin Aeroportul Heathrow.

Aşa cum vom vedea în continuare, algoritmii de recunoaştere facială sunt, cu siguranţă, mai buni decât oamenii. Însă, din moment ce sunt folosiţi pentru a vâna infractorii, iar consecinţele identificării greşite sunt atât de grave, utilizarea lor ridică o întrebare importantă: cât de

uşor ar putea fi confundată o persoană cu alta? câţi dintre noi au vreo sosie pitită pe undeva, asemenea lui Steve Talley?

Un studiu din 2015 părea să sugereze că şansele să ai o dublură în viaţa reală (indiferent dacă este un spărgător de bănci sau orice altceva) sunt incredibil de mici. Teghan Lucas, de la Universitatea din Adelaide, a făcut, cu migală, câte opt măsurători faciale pe fotografiile a 4.000 de persoane şi nu a reuşit să găsească nici o singură potrivire printre ele, ceea ce a făcut-o să tragă concluzia că şansele ca două persoane să aibă exact aceeaşi faţă erau mai mici de una la un trilion.53 în conformitate cu aceste calcule, Talley fusese îngrozitor de ghinionist Ţinând cont de faptul că geamănul său malefic locuia în apropiere şi că întâmplarea făcea ca acesta să fie un infractor, ne-am putea aştepta să treacă încă zeci de mii de ani înainte ca alt suflet amărât să cadă victimă aceleiaşi experienţe nefericite.

Avem, însă, motive să bănuim că toate numerele acestea nu prea se potrivesc. Deşi este destul de dificil să-ţi imaginezi că te-ai putea întâlni cu cineva care are exact chipul tău, dovezile existenţei unor persoane neînrudite, care nu se cunosc, dar care arată identic, par să fie mai des întâlnite decât ar sugera cercetarea lui Lucas.

Să-l luăm pe Neil Douglas, de exemplu, care se îmbarca într-un avion cu destinaţia Irlanda când şi-a dat seama că dublura lui era aşezată pe scaunul său! Selfie-ul pe care şi l-au făcut, cu toţi pasagerii din avion râzând la unison în fundal, a devenit repede viral şi, în scurt timp, bărbaţi roşcaţi şi cu barbă din întreaga lume îşi trimiteau propriile poze, să demonstreze că şi ei aveau o înfăţişare asemănătoare. „Cred chiar că la un moment dat formam o mică armată”, a declarat Neil pentru postul de televiziune BBC.54

Chiar şi eu am o istorie personală de adăugat la toate celelalte. Când aveam 22 de ani, un prieten mi-a arătat o fotografie pe care o văzuse pe pagina de My Space a unei formaţii muzicale din partea locului. Era un colaj de fotografii făcute în timpul unui concert la care nu participasem, în care apăreau câteva persoane care se distrau, iar una dintre acestea părea ciudat de familiară. Doar ca să mă asigur că nu-mi pierdusem, fără să ştiu, cunoştinţa într-o noapte şi mă rătăcisem la o petrecere la care nu-mi aminteam să fi participat, i-am trimis un e-mail solistului formaţiei, care a confirmat ceea ce bănuiam şi eu: sosia mea iubitoare de synthpop avea o viaţă socială mai activă decât a mea.

Aşadar, iată-ne, pe Talley, Douglas şi pe mine, având fiecare cel puţin câte o sosie, dacă nu chiar mai multe. Suntem trei inşi dintr-o populaţie de şapte miliarde şi jumătate de oameni şi suntem deja cu mult sub estimarea lui Lucas, de un caz la un trilion.

Există un motiv pentru această discrepanţă. Totul se reduce la definirea de către cercetători a noţiunii de „identic”. Studiul lui Lucas cerea ca măsurătorile a două persoane să coincidă perfect. Deşi Neil şi sosia lui sunt incredibil de asemănători, dacă o nară sau lobul unei urechi s-ar deosebi fie şi numai cu un milimetru, ei nu ar mai fi consideraţi sosii 100% prin prisma criteriilor stabilite de cercetătoare.

Dar chiar şi atunci când compari două imagini ale aceleiaşi persoane, măsurătorile exacte nu vor reflecta felul în care fiecare dintre noi ne schimbăm în permanenţă prin înaintarea în vârstă, boală, oboseală, expresiile pe care le adoptăm sau felul în care chipurile ne sunt distorsionate de unghiul camerei video. Dacă se încearcă surprinderea particularităţilor unei feţe doar prin măsurarea acesteia în milimetri, se va descoperi tot atâta variaţie în chipul unei singure persoane, ca şi între două persoane diferite. Mai simplu spus: măsurătorile nu sunt suficiente pentru a distinge un chip de altul.

Deşi s-ar putea să nu fie perfect identici, îmi pot imagina fără dificultate că i-aş putea confunda pe Neil şi pe geamănul său necunoscut în acea fotografie. Ca şi în cazul lui Talley – bietul Steve nici măcar nu semăna *chiar* atât de mult cu adevăratul infractor şi, cu toate acestea, imaginile au fost interpretate greşit de specialiştii FBI, aşa că el a fost acuzat de o infracţiune pe care nu o comisese, fiind aruncat într-o celulă de maximă securitate.

Aşa cum au demonstrat ofiţerii de la paşapoarte, este uimitor de uşor să confunzi chipuri care nu-ţi sunt familiare, chiar şi atunci când acestea sunt doar vag asemănătoare, iar asta înseamnă că noi, oamenii, suntem surprinzător de incapabili să recunoaştem, să identificăm, persoanele necunoscute. Este şi motivul pentru care o prietenă de-ale mele susţinea că abia a reuşit să reziste până la sfârşitul frumosului film al lui Cristopher Noian, *Dunkirk*, pentru că se chinuia să deosebească actorii între ei. De aceea consideră adolescenţii util» să împrumute” cartea de identitate a unui prieten cu câţiva ani mai mare, ca să cumpere băuturi alcoolice. Şi acesta este motivul pentru care Innocence Project, un ONG din SUA, estimează că identificarea greşită de către martorii oculari joacă un rol semnificativ în peste 70% dintre condamnările pe nedrept.55

Cu toate acestea, deşi un martor ocular ar putea să-l confunde cu uşurinţă pe Neil cu tovarăşul lui de călătorie,

mama sa nu ar avea, cu siguranţă, nicio dificultate în a-şi identifica fiul în acea fotografie. Când este vorba despre persoane cunoscute, ne pricepem extrem de bine să le recunoaştem fizionomiile, chiar şi atunci când este vorba despre sosiile din viaţa reală: poţi confunda cu uşurinţă doi gemeni identici dacă nu-ţi sunt decât simple cunoştinţe, însă îi vei putea deosebi fără dificultate odată ce ajungi să îi cunoşti mai bine.

Iar aici ajungem într-un punct critic:» nu-s asemănătoare lucrurile similare, ci-s asemănătoarea acelea care-mi par mie" astfel, ca să parafrazez un proverb bine-cunoscut. În lipsa unei definiţii foarte clare a asemănării, nu poţi să măsori *cât* de diferite sunt două feţe şi nu există nicio valoare în funcţie de care să judecăm pentru a afirma indubitabil că două chipuri sunt identice. Nu poţi să defineşti ce înseamnă să fii o sosie sau să spui cât de comună este o anumită faţă şi nici (ceea ce este extrem de important) nu poţi să stabileşti care este probabilitatea ca în două imagini să apară acelaşi individ.

Asta înseamnă că recunoaşterea facială, ca metodă de identificare, nu este asemenea ADN-ului, care stă semeţ pe un fundament statistic solid. Când în criminalistică se foloseşte testarea ADN-ului, profilarea se concentrează asupra anumitor bucăţi din genom despre care se ştie că variază foarte mult de la un individ la altul. Amploarea acestei variaţii ne dă răspunsul: dacă secvenţa de ADN dintr-un eşantion de ţesut uman găsit la locul unei infracţiuni se potriveşte cu secvenţa din proba de salivă a unui suspect, înseamnă că poţi să calculezi probabilitatea ca ambele să provină de la aceeaşi persoană. Mai înseamnă şi că poţi să indici cu exactitate şi care sunt şansele ca o fiinţă ghinionistă să aibă, întâmplător, o secvenţă de ADN identică sub aceleaşi aspecte.56 Cu cât foloseşti mai mulţi markeri, cu atât sunt mai reduse şansele unei identificări greşite; astfel, alegând numărul de markeri pe care să-i testeze, orice sistem judiciar din lume are deplina putere de a decide asupra nivelului de

dubiu pe care este dispus să-l tolereze.57

Chiar dacă feţele noastre par atât de intrinsec legate de cine suntem, fără să cunoaştem variaţia care se manifestă între indivizi, practica identificării infractorilor după chipurile lor nu are o bază ştiinţifică riguroasă. Când se pune problema identificării persoanelor după fotografii – ca să citǎm o prezentare făcută de o echipă de criminalişti ai FBI: „Lipsa statisticilor înseamnă că, de fapt, concluziile se bazează, în ultimă instanţă, pe opinii personale”.58

Din nefericire, utilizarea algoritmilor pentru recunoaşterea facială nu soluţionează această dilemă morală, ceea ce este un motiv foarte bun să manifestăm prudenţă atunci când îi folosim pentru identificarea infractorilor. Asemănarea şi identitatea nu sunt acelaşi lucru şi nici nu vor fi vreodată, oricât de precis ar deveni algoritmul.

Şi mai există un motiv întemeiat pentru a proceda cu atenţie atunci când folosim algoritmi de recunoaştere facială. Ei nu sunt chiar atât de buni la recunoaşterea chipurilor cum s-ar putea crede.

Algoritmii funcţionează folosind una din cele două abordări principale. Primul tip de algoritm construieşte un model 3D al unei feţe fie prin combinarea unei serii de imagini 2D, fie prin scanarea cu ajutorul unei camere video speciale, cu infraroşu. Este metoda adoptată de sistemul Face ID, pe care Apple îl foloseşte la iPhone-urile sale. Algoritmii aceştia au găsit o cale să ocolească problemele puse de diferitele expresii faciale şi de înaintarea în vârstă, concentrându-se asupra zonelor feţei care au ţesut rigid şi formaţiuni osoase, cum ar fi curbura orbitei ochiului sau septul nazal.

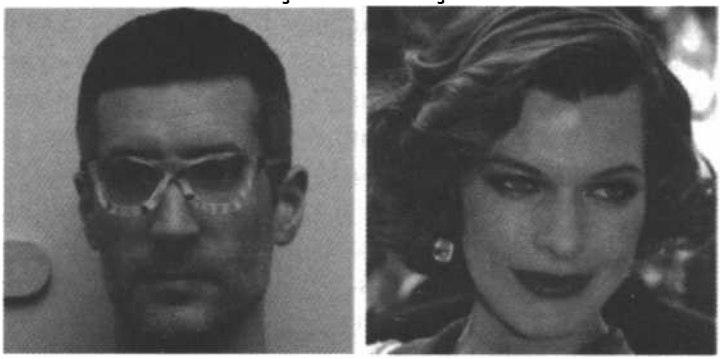
Apple a susţinut că şansele ca o persoană oarecare să-ţi deblocheze telefonul care utilizează Face ID sunt de una la un milion, însă algoritmul nu funcţionează absolut impecabil. El poate fi păcălit de gemeni59, de fraţi sau surori60 şi de către copiii care şterpelesc telefoanelor

părinţilor lor. (Curând după lansarea Face ID, pe internet a apărut o înregistrare video cu un băiat de zece ani care a reuşit să păcălească sistemul de recunoaştere facială de pe iPhone-ul mamei sale. Acum aceasta îşi şterge SMS-urile dacă în ele apare ceva ce nu ar dori să fie văzut de către fiul ei.) 61 S-au mai relatat cazuri în care algoritmul a fost păcălit de o mască specială, printată la o imprimantă 3D, cu nişte imagini infraroşii lipite în locul ochilor.62 Toate acestea înseamnă că, deşi algoritmul ar putea fi destul de bun ca să-ţi blocheze telefonul, probabil că nu este suficient de performant ca să fie utilizat pentru a-ţi permite accesul la conturile bancare.

Aceşti algoritmi 3D nu sunt de prea mare folos nici la scanarea fotografiilor de paşaport sau a înregistrărilor camerelor CCTV. Pentru asta ai nevoie de un al doilea tip de algoritm, care se limitează la imaginile 2D şi foloseşte o abordare statistică. Aceşti algoritmi nu au în vedere neapărat reperele pe care dumneavoastră sau eu le-am putea recunoaşte ca trăsături distinctive, ci construiesc o descriere statistică a tiparelor cu zone luminoase şi întunecate dintr-o imagine. Asemenea algoritmilor din capitolul „Medicina”, construiţi să recunoască un câine într-o imagine, cercetătorii şi-au dat de curând seama că, în loc să te bazezi pe oameni care să decidă ce tipare vor funcţiona mai bine, poţi să faci algoritmul să înveţe care sunt cele mai bune combinaţii prin metoda încercării şi a erorii, aplicată la o imensă bază de date cu feţe umane. De regulă, acest lucru se realizează utilizând reţele neurale. Cu genul acesta de algoritmi s-au făcut, recent, progrese importante în materie de performanţă şi acurateţe. Această performanţă, însă, are şi un preţ. Nu este întotdeauna clar *cum* decide algoritmul dacă o faţă este asemenea alteia.

Înseamnă, totodată, şi că aceşti algoritmi de ultimă generaţie pot fi relativ uşor păcăliţi. Din moment ce funcţionează prin detectarea unei descrieri statistice a tiparelor de lumină şi umbră de pe o faţă, îi poţi păcăli

chiar şi prin purtarea unor ochelari mai zbanghii pe care e imprimat un model aiurit şi care induce în eroare. Şi mai bine încă: prin crearea unui model năucitor specific menit să indice chipul altei persoane poţi, efectiv, să faci algoritmul să creadă că eşti acea persoană – aşa cum procedează bărbatul din imaginea următoare, care poartă nişte ochelari ce îl fac să „arate ca „actriţa Milla Jovovich.63 Adică poţi să foloseşti ochelarii să te deghizezi? Clark Kent ştia el ce ştia!



Dar, făcând abstracţie de atacurile personalizate cu ochelari caraghioşi, abilităţile de recunoaştere ale acestor algoritmi statistici au stârnit numeroase titluri admirative în mass-media, precum cele care au întâmpinat apariţia Face Net al lui Google. Pentru a-i pune la încercare abilităţile de recunoaştere, lui Face Net i s-a cerut să identifice 5.000 de imagini cu chipurile unor celebrităţi. Persoane specializate în acest tip de recunoaşte încercaseră anterior să îndeplinească aceeaşi sarcină şi se descurcaseră excepţional de bine, obţinând un procent de 97, 5% de identificări corecte (deloc surprinzător, din moment ce chipurile acestor oameni celebri le erau deja familiare participanţilor).64 Face Net s-a descurcat chiar mai bine, obţinând procentul fenomenal de 99, 6% de identificări corecte.

La prima vedere, s-ar părea că maşinile acestea deţin nişte capacităţi supraumane de recunoaştere a feţelor. Asta lasă impresia unui rezultat de excepţie, fără îndoială, suficient de bun pentru a justifica utilizarea algoritmilor

pentru identificarea infractorilor. Numai că aici se ascunde o capcană. Cinci mii de feţe înseamnă, în realitate, o cifră deplorabil de mică pe care să-ţi testezi algoritmul. Dacă acesta urmează să fie pus la lucru pentru combaterea infracţionalităţii, va trebui să găsească o faţă din milioane, nu din mii de alte feţe.

Asta pentru că poliţia britanică deţine acum o bază de date de 19 milioane de imagini cu feţele noastre, creată din toate acele fotografii ale indivizilor arestaţi în baza suspiciunii de a fi comis o infracţiune. Totodată, FBI are o bază de date de 411 milioane de imagini în care se afirmă că ar apărea jumătate din toţi americanii adulţi.65 Iar în China, unde baza de date a cărţilor de identitate emise oferă un acces facil la miliarde de feţe, autorităţile au investit deja masiv în recunoaşterea facială. Există camere de supraveghere instalate pe străzi, în metrouri şi aeroporturi, care se presupune că vor detecta orice, de la infractori căutaţi de poliţie până la pietoni care traversează strada prin locuri nepermise, prin diverse oraşe din ţară.66 (Există chiar o sugestie ca şi abaterile minore ale cetăţenilor din lumea reală, cum ar fi aruncarea gunoiului pe jos, să fie incluse în sistemul de punctaj Sesame Credit, ceea ar atrage asupra acestora toate pedepsele pe care le-am dezvăluit la capitolul „Datele personale informatizate”.)

Iată unde este problema: şansele de identificare greşită se înmulţesc spectaculos odată cu numărul feţelor adăugate. Cu cât sunt mai numeroase feţele printre care caută algoritmul, cu atât cresc şansele ca acesta să găsească două feţe care să arate la fel. Prin urmare, în momentul în care încerci să foloseşti aceiaşi algoritmi în cataloage care conţin tot mai multe feţe, precizia lor scade.

Ar fi ca şi când mi-aţi cere să potrivesc ce carte de identitate îi corespunde fiecăreia dintre zece persoane necunoscute, iar atunci când am obţinut punctajul maxim să susţineţi că am reuşit să identific corect feţele în 100%

din cazuri, după care să mă lăsaţi să mă plimb la întâmplare prin centrul New York-ului, să identific infractorii recunoscuţi. În mod inevitabil, acurateţea mea va scădea.

Acelaşi lucru se întâmplă şi cu algoritmii. În anul 2015, Universitatea din Washington a organizat aşa-numita competiţie Mega Face, în care oameni din toată lumea au fost invitaţi să îşi testeze algoritmii de recunoaştere facială pe o bază de date de un milion de feţe.67 Deşi aceasta bază de date se apropie, ca dimensiune, de cataloagele deţinute de unele autorităţi guvernamentale, ea tot este considerabil mai mică decât aceştia. Chiar şi aşa, algoritmii nu s-au descurcat prea bine în această competiţie.

Face Net al companiei Google – care funcţionase aproape perfect în cazul celebrităţilor, dintr-odată, a obţinut doar 75%[77] ca rată de identificare.68 Alţi algoritmi au ajuns la o rată de succes de-a dreptul jalnică, de numai 10%. În momentul în care scriu această carte, cea mai bună ofertă din lume este aceea a chinezilor, un algoritm numit Tencent Youâu Lab, care reuşeşte să obţină o rată de recunoaştere de 83, 29%.69

Altfel spus, judecând după aceste procente, dacă te-ai afla în căutarea unui anumit infractor într-o listă digitală de milioane de feţe, în cel mai bun caz, în unul din şase cazuri nu vei găsi persoana potrivită.

Ar trebui să mai adaug că acest domeniu progresează foarte repede. Ratele de identificare corectă sunt în creştere uşoară şi nimeni nu poate să spună, cu certitudine, ce se va întâmpla în lunile sau în anii care vor urma. Ceea ce pot să afirm este că diferenţele de luminozitate, postură, calitate a imaginii şi aspect general fac ca recunoaşterea cu acurateţe şi fiabilitate a feţelor să

77 Această cifră se referă la cât de multe potriviri reale de chipuri a omis algoritmul atunci când a fost setat să evite identificările greşite, în capitolul „Justiţia” sunt descrise mai multe tipuri de erori pe care le poate face un algoritm, iar diferitele modalităţi de măsurare a acurateţei sunt prezentate în capitolul „Medicina” (n. a.).

fie o problemă într-adevăr foarte complicată. Suntem oarecum departe de a obţine o precizie perfectă în identificare atunci când căutările se fac într-o bază de date de 411 milioane de feţe sau de a reuşi să găsim acea potrivire de unu la un trilion în cazul sosiilor.

STABILIREA UNUI ECHILIBRU

Acestea sunt realităţi, dar nu neapărat şi motive de renunţare. Există algoritmi destul de buni pentru a fi folosiţi în anumite situaţii. În Ontario, de exemplu, persoanele care suferă de dependenţă de jocurile de noroc se pot înscrie voluntar pe o listă în virtutea căreia li se interzice accesul în cazinouri. Dacă hotărârea lor de a rămâne pe drumul cel drept nu mai e fermă, chipul le va fi descoperit de algoritmii de recunoaştere ai cazinourilor, care îi vor alerta pe angajaţi să le ceară în mod politicos pariorilor să părăsească incinta.70 Sistemul este, cu siguranţă, nedrept faţă de toţi cei care sunt împiedicaţi din greşeală să petreacă o noapte la jocul de ruletă, însă cred că acest preţ merită plătit dacă asta înseamnă să ajuţi o persoană dependentă de jocuri de noroc, aflată în plin proces de renunţare la viciu, să reziste tentaţiei de a-şi relua vechile apucături.

Acelaşi lucru este valabil şi în cazul vânzării cu amănuntul. Personalul de pază din magazine obişnuia să tapeteze pereţii birourilor cu fotografii Polaroid ale hoţilor din magazine; în prezent, algoritmii pot să-ţi compare faţa cu acelea din baza de date cu hoţii cunoscuţi chiar din clipa în care ai trecut pragul magazinului. Dacă faţa ta se potriveşte cu aceea a unui infractor bine-cunoscut, este trimisă o alertă pe smartphone paznicilor care te pot găsi apoi prin raioane.

Magazinele au o motivaţie puternică să îşi dorească utilizarea acestui gen de tehnologie. Aproximativ 3, 6 milioane de infracţiuni sunt comise anual în magazine numai în Marea Britanie, costându-i pe comercianţi uluitoarea sumă de 660 de milioane de lire sterline.71 Când te gândeşti că în anul 2016 au fost omorâţi violent 91

de oameni suspectaţi că ar fi furat din magazine din SUA72, acesta este un argument în favoarea faptului că o metodă de a-i împiedica pe infractorii insistenţi să intre într-un magazin înainte ca situaţia să escaladeze ar fi benefică pentru toată lumea.

Numai că această soluţie high-tech la furtul din magazine are şi dezavantaje: în primul rând, confidenţialitatea (Face First, unul dintre furnizorii importanţi al acestui tip de software de securitate, susţine că nu stochează imaginile clienţilor obişnuiţi, însă magazinele folosesc, cu siguranţă, recunoaşterea facială pentru ne urmări şi a ne analiza comportamentul la shopping). Iar apoi se mai pune problema: cine ajunge să figureze pe lista neagră digitală. De unde ştim că toţi cei de pe listă se află acolo din motive întemeiate? Cum rămâne cu prezumţia de nevinovăţie? Ce se întâmplă cu cei care ajung pe listă în mod accidental: cum vor reuşi să iasă de acolo? La toate acestea se adaugă şi faptul că apare posibilitatea identificării greşite de către un algoritm, care nu poate fi niciodată absolut precis.

Întrebarea care se pune este dacă avantajele depăşesc dezavantajele. Răspunsul nu este simplu. Nici măcar comercianţii nu sunt, în unanimitate, de aceeaşi părere. Unii adoptă cu entuziasm această tehnologie, pe când alţii o resping – inclusiv Walmart, care a pus capăt testării algoritmului Face First în magazinele sale după ce nu a dat randamentul pe care-l spera compania în urma investiţiei în acesta.73

În cazul infracţionalităţii, însă, echilibrul dintre rău şi bine pare mult mai clar delimitat. Este adevărat, aceşti algoritmi nu sunt singurii utilizaţi, căci fundamentările lor statistice sunt cam precare. Amprentarea nu are nicio rată de eroare cunoscută74, nici analiza urmelor de muşcătură, nici tiparul stropilor de sânge75 şi nici balistica76. De fapt, dacă ar fi să dăm crezare unei lucrări din 2009 a Academiei Naţionale de Ştiinţe a SUA, niciuna dintre tehnicile de investigaţie criminalistică, exceptând testarea

ADN, nu poate „să demonstreze existenţa unei legături între o probă şi un anumit individ sau o anumită sursă”.77 Totuşi, nimeni nu poate să nege că toate s-au dovedit nişte instrumente extraordinar de utile în munca de poliţie – atât timp cât nimeni nu se bazează foarte mult pe probele pe care le generează. Însă ratele de acurateţe chiar şi ale celor mai sofisticaţi algoritmi de recunoaştere facială lasă mult de dorit. Se susţine că, dacă există chiar şi cel mai mic risc să apară mai multe cazuri precum acela al lui Steve Talley, atunci o tehnologie imperfectă nu ar trebui utilizată pentru a contribui la răpirea libertăţii unei persoane. Singura problemă este că poveşti ca aceea a lui Talley nu prea ajung să descrie întregul tablou. Pentru că, deşi utilizarea recunoaşterii faciale pentru prinderea infractorilor are dezavantaje uriaşe, ea are şi avantaje imense.

UN COMPROMIS DELICAT

În mai 2015, un bărbat fugea pe străzile din Manhattan atacând la întâmplare trecătorii cu un ciocan mare şi negru.

La început, s-a îndreptat în fugă către un grup de oameni de lângă Empire State Building şi l-a lovit cu putere în ceafa pe un tânăr de 20 de ani. Şase ore mai târziu, s-a îndreptat spre sud, spre Union Square, şi cu acelaşi ciocan a atacat o femeie care stătea liniştită pe o bancă din parc, lovind-o în tâmplă. Doar câteva minute mai târziu, şi-a făcut din nou apariţia, vizând de data asta o femeie de 33 de ani care mergea pe o stradă care mărginea parcul.78 Folosind înregistrările camerelor de supraveghere care surprinseseră atacurile, un algoritm de recunoaştere facială a reuşit să-l identifice pe David Baril, un bărbat care postase pe Instagram, cu câteva luni înaintea atacurilor, o poză cu un ciocan de pe care picura sânge.79 Acesta s-a declarat vinovat de acuzaţiile care rezultau din atacuri şi a fost condamnat la 22 de ani de închisoare.

Şi cazurile vechi sunt readuse la viaţă de progresele

recunoaşterii faciale. În 2014, un algoritm a adus în faţa justiţiei un bărbat de cetăţenie americană care trăia de 15 ani ca un fugar, sub un nume fals. Neil Stammer se sustrăsese de la executarea condamnării sale la închisoare în timpul eliberării pe cauţiune pentru acuzaţii care includeau abuz sexual asupra minorilor şi răpire; a fost arestat din nou atunci când poza lui de persoană căutată de FBI a fost comparată din nou cu baza de date a paşapoartelor şi a fost găsită o potrivire cu o persoană care locuia în Nepal, a cărei fotografie de paşaport purta un nume diferit.80

După vara anului 2017, când au murit opt persoane în urma unui atac terorist pe London Bridge, îmi dau seama cât de util poate fi un sistem care ar folosi un astfel de algoritm. Youssef Zaghba a fost unul dintre cei trei bărbaţi care au intrat cu maşina în pietoni, înainte de a porni să înjunghie oameni la întâmplare în apropiere, în Borough Market. Figura pe lista persoanelor suspectate de activităţi teroriste din Italia şi ar fi putut fi identificat automat de un algoritm de recunoaştere facială înainte de a intra în Marea Britanie.

Dar cum decizi să faci acel compromis între confidenţialitate şi protecţie, între echitate faţă de toată lumea şi siguranţă? Cât de mulţi Steve Talley suntem dispuşi să acceptăm în schimbul identificării rapide a unora precum David Baril şi Youssef Zaghba?

Să aruncăm o privire asupra statisticilor oferite de NYPD. În anul 2015, acest departament de poliţie a raportat identificarea cu succes a 1 700 de suspecţi, ceea ce a dus la 900 de arestări, dar, totodată, a identificat greşit cinci persoane.81 Oricât ar fi de îngrijorătoare eroarea făcută cu fiecare dintre cele cinci, rămâne întrebarea: este acesta un procent acceptabil? Este un preţ pe care suntem dispuşi să-l plătim pentru a reduce rata criminalităţii?

După cum se pare, algoritmii lipsiţi de dezavantaje precum geo-profilarea lui Kim Rossmo, despre care am

discutat la începutul capitolului, reprezintă, mai degrabă, excepţia decât regula. Atunci când se ajunge la combaterea infracţionalităţii, oriunde ai căuta, vei găsi algoritmi foarte promiţători sub un anumit aspect, dar care pot fi foarte îngrijorători sub un altul. Pred Pol, Hunchlab, Strategic Subiect List şi recunoaşterea facială – toate promit să ne rezolve problemele, însă toate creează altele noi pe parcurs.

După părerea mea, necesitatea reglementării urgente a utilizării algoritmilor nu a fost niciodată mai acută decât în cazul infracţionalităţii, unde simpla existenţă a acestor sisteme informatice ridică întrebări serioase pentru care nu există răspunsuri facile. Va trebui să ne confruntăm cumva cu aceste dileme dificile. Ar fi bine să acceptăm numai algoritmii pe care îi putem înţelege sau înăuntrul cărora putem arunca o privire, ştiind că a-i lua de sub oblăduirea proprietarilor lor ar putea duce la pierderea eficienţei (iar rata infracţionalităţii va creşte)? Sau să respingem din start orice sistem matematic care are incluse din construcţie unele scăpări sau care s-a dovedit capabil de eroare, ştiind că în felul acesta vom impune algoritmilor un standard mai înalt decât îi cerem societăţii în care trăim? Şi cât de tendenţios înseamnă „prea tendenţios”? Când să acordăm prioritate victimelor unor infracţiuni care puteau fi prevăzute, în defavoarea victimelor pe care le face algoritmul?

Într-o anumită măsură, asta se reduce la a decide, ca societate, ce credem noi că ar însemna succesul. Care este prioritatea noastră? Să menţinem infracţionalitatea la cel mai scăzut nivel cu putinţă? Sau să protejăm, mai presus de orice, libertatea celor nevinovaţi? Cât de mult dintr-una aţi fi dispuşi să sacrificaţi de dragul celeilalte?

Gary Marx, profesor de sociologie la MIT, a descris foarte bine această dilemă într-un interviu pe care l-a dat revistei *The Guardian*, în care afirma:

Uniunea Sovietică avea un nivel remarcabil de scăzut de delicte minore de stradă în perioada în care controlul

totalitar şi autoritar era cel mai atroce. Dar, Dumnezeule, cu ce preţ obţinuse asta? 82

Este posibil să decidem, până la urmă, că ar trebui să le impunem algoritmilor anumite limite până la care să aibă acces, pentru ca anumite lucruri să nu fie analizate şi calculate. Această perspectivă s-ar putea aplica şi dincolo de lumea infracţionalităţii. Nu neapărat din cauza lipsei de străduanie din partea algoritmilor înşişi, ci poate – doar „poate" – pentru că există şi lucruri care depăşesc competenţa unei maşini lipsite de simţire.

Capitolul 7

ARTA

Justin căzuse pe gânduri. Pe 4 februarie 2018, în sufrageria locuinţei lui din Memphis, statul Tennessee, urmărea Super Bowl, meciul de super-cupă de fotbal american şi mânca bombonele M&M. Mai devreme, în aceeaşi săptămână, sărbătorise împlinirea a 37 de ani, iar acum – ceva ce devenise pentru el o tradiţie anuală – cugeta la ce se alesese de viaţa lui.

Ştia că ar trebui să fie recunoscător, chiar ştia asta. Avea o viaţă cât se poate de confortabilă. O slujbă de birou de la nouă la cinci, un acoperiş deasupra capului şi o familie iubitoare. Numai că îşi dorise dintotdeauna ceva mai mult. În copilărie şi adolescenţă, fusese tot timpul convins că era menit să aibă faimă şi avere.

Cum ajunsese, atunci, să fie atât de... normal? „Totul mi s-a tras de la trupa aceea de băieţi", se gândea el, aceea în care intrase la 14 ani. „Dacă am fi avut succes, acum totul ar fi fost diferit". Dar din cine ştie ce motiv, trupa fusese un fiasco. Succesul nu se prea arătase pentru bietul, sărmanul Justin Timberlake.

Abătut, mai deschise o bere şi îşi imagină cum ar fi putut să fie viaţa sa. Pe ecranul televizorului, reclamele de la Super Bowl luară sfârşit. Începu muzica, anunţând marele spectacol al reprizei a doua. Iar într-un univers paralel – practic identic cu acesta cu excepţia unui singur detaliu – un alt Justin Timberlake de 37 de ani, din

Memphis, intră în scenă.

MAI MULTE LUMI

De ce adevăratul Justin Timberlake avea atât de mult succes? Şi de ce gânditorul Justin Timberlake eşuase? Unii (inclusiv eu, cea care am fost la vârsta de 14 ani) [78] ar putea susţine că succesul lui Justin este binemeritat: înzestrarea nativă, aspectul fizic plăcut, talentul de dansator şi valoarea artistică a muzicii sale au făcut din celebritate o certitudine. Alţii, însă, ar putea să nu fie de acord. Poate vor susţine că Timberlake nu are nimic deosebit, ca, de altfel, oricare alt superstar al muzicii pop adulat de legiuni întregi de fani. Este uşor să găseşti oameni talentaţi care se pricep să cânte şi să danseze – starurile nu sunt decât cei care au avut noroc.

Bineînţeles că nimeni nu poate fi sigur de asta. Nu fără să construiască un număr de lumi identice paralele, aducându-l pe Timberlake în fiecare dintre ele şi urmărind evoluţia tuturor întrupărilor sale ca să vadă dacă are succes de fiecare dată. Din păcate, crearea unui multivers artificial ne depăşeşte pe mulţi dintre noi, însă, dacă ne situăm sub nivelul lui Timberlake şi luăm în considerare muzicienii mai puţin cunoscuţi, se poate studia rolul relativ pe care îl joacă şansa şi talentul în popularitatea unui hit.

Exact asta a fost ideea care a stat la baza unui experiment celebru, realizat de cercetătorii Matthew Salganik, Peter Dodds şi Duncan Watts în 2006, în care au fost create mai multe lumi digitale.1 Oamenii de ştiinţă şi-au construit propriul player audio on-line, un fel de versiune foarte primitivă a Spotify, şi au filtrat vizitatorii spre opt website-uri paralele de muzică, pe fiecare din ele fiind instalate aceleaşi 48 de cântece ale unor artişti necunoscuţi.

În ceea ce a devenit cunoscut drept *the Music* Lob2, 14 341 de pasionaţi de muzică au fost invitaţi să se autentifice în player, să asculte clipurile de pe fiecare

78 Pe de altă parte, eul meu de 14 ani era un mare fan al duoului PJ & Duncan, aşa că nu prea avea habar de multe! (n. a.).

playlist, să evalueze cântecele şi să descarce melodia care le plăcea mai mult.

La fel ca în Spotify-ul real, vizitatorii puteau vedea dintr-o privire ce muzică ascultau alte persoane în „lumea” lor. Alături de numele artistului şi de titlul cântecului, participanţii vedeau de câte ori fusese descărcată piesa respectivă în „lumea” lor. Toate contoarele au început de la zero, iar în timp, pe măsură ce cifrele se schimbau, cântecele cele mai apreciate în fiecare dintre cele opt topuri paralele începeau să se detaşeze cu claritate.

Între timp, pentru a obţine nivelul „realei” popularităţi a cântecelor, echipa de cercetători a construit şi o „lume de control”, în care preferinţele vizitatorilor nu puteau fi influenţate de acelea ale altor utilizatori. Aici, cântecele aveau să apară pe pagină într-o ordine aleatorie – într-o grilă sau pe o listă – însă statisticile descărcărilor urmau să fie ascunse vederii.

Rezultatele obţinute i-au intrigat pe toţi. Toate „lumile” au fost de acord că anumite cântece erau de-a dreptul jalnice. Altele erau veritabile hituri: acestea au sfârşit prin a deveni populare în fiecare lume, chiar şi în aceea în care vizitatorii nu puteau să vadă numărul descărcărilor. Dar între hiturile incontestabile şi chixurile absolute, artiştii se puteau situa cam la orice nivel de succes.

Să luăm, de exemplu, 52 Metro, o formaţie punk din Milwaukee al cărei cântec, „Lockdown”, a fost extrem de popular într-una dintre lumi, unde a ajuns, în final, pe primul loc al topului, dar care, într-o altă lume, a dat chix, situându-se abia pe locul al patruzeci şi optulea. Exact acelaşi cântec, la concurenţă cu aceeaşi listă de cântece; doar că în lumea aceasta 52 Metro nu a prins deloc.3 Se părea că, uneori, succesul chiar e o chestiune de noroc.

Deşi drumul spre primul loc nu era bătut în cuie, cercetătorii au descoperit că era mult mai probabil ca vizitatorii să descarce melodiile despre care ştiau că fuseseră apreciate de către alţi ascultători. Dacă un cântec

mediocru ajungea, din întâmplare, în topul preferinţelor mai devreme, popularitatea lui putea să crească asemenea unui adevărat bulgăre de zăpadă. Mai multe descărcări generau încă şi mai multe descărcări. Popularitatea percepută devenea o popularitate reală, aşa încât eventualul succes nu era decât o întâmplare care ajunsese să se amplifice în timp.

Aceste rezultate aveau o motivare: fenomenul cunoscut de către psihologi drept „dovadă socială”[79] sau „influenţare prin intermediul informaţiilor primite din societate”. Ori de câte ori nu dispunem de informaţii suficiente pentru a lua noi înşine decizii, obişnuim să imităm comportamentul celor din jur. De aceea în sălile de spectacol sunt introduse, în secret, printre spectatori, câteva persoane care să aplaude şi să ovaţioneze la momentul potrivit. Imediat ce-i auzim pe alţii aplaudând, este foarte probabil să o facem şi noi. Când vine vorba despre alegerea muzicii, nu înseamnă neapărat că preferăm să ascultăm aceleaşi cântece ca şi alte persoane, ci doar că popularitatea lor este o cale rapidă de a ne feri de dezamăgire. Sau, după cum a declarat Salganik însuşi revistei *Live Science*:

Oamenii sunt puşi în faţa prea multor opţiuni. Din moment ce nu poţi să le asculţi pe toate, o scurtătură firească este aceea de a asculta ce ascultă şi alte persoane.4

Folosim popularitatea ca indicator al calităţii în toate formele de divertisment. Un studiu din 2007, de exemplu, a cercetat impactul pe care îl are asupra percepţiei publicului despre o anumită carte faptul că aceasta a apărut în topul bestsellerurilor din publicaţia *New York Times*. Exploatând particularităţile modului de alcătuire a acestui clasament, Alan Sørensen, autorul studiului, a urmărit cărţile de succes care ar fi trebuit incluse, în

79 Social proo/(engl.), termen inventat de prof. univ. Robert Cialdini, cunoscut în limba română şi sunt denumirea peiorativă „efect de turmă” (n. red.).

funcţie de numărul exemplarelor vândute, însă care – din cauza întârzierilor şi omisiunilor accidentale – nu mai apăruseră pe listă şi le-a comparat cu acelea care reuşiseră să figureze. A descoperit un efect dramatic: simplul fapt că unele cărţi figurau pe listă dusese la o creştere a vânzării lor, în medie, cu 13 – 14% şi la o creştere a vânzărilor cu 57% în cazul autorilor aflaţi la debut.

Cu cât folosim mai multe platforme pentru a vedea ce este popular într-un anumit moment (liste de bestselleruri, clasamentele de pe Amazon, punctajul obţinut de un film pe Rotten Tomatoes, topurile hiturilor de pe Spotify), cu atât dovada socială va avea un impact mai mare. Efectul este amplificat şi mai mult atunci când suntem bombardaţi cu milioane de opţiuni, însoţite de campaniile lor de marketing, impactul celebrităţii creatorului, de publicitatea făcută în mass-media sau cu ajutorul aprecierilor critice şi când absolut toate ne reclamă atenţia.

Toate acestea înseamnă că, uneori, un cântec penibil poate ajunge în topul preferinţelor (şi nu spun asta pentru c-aş fi cinică). Prin anii 1990 se zvonea că producătorii britanici de muzică – în deplină cunoştinţă de cauză – încheiaseră un pariu: care dintre ei reuşea să aducă în topurile muzicale cel mai prost cântec cu putinţă. În urma pariului se spune că a fost aleasă o formaţie de fete numită Vanilia, al cărei cântec de debut, „No way, no way, mă, na, mă, na” se inspira din cunoscutul cântecel al celebrelor păpuşi Muppets. Formaţia aceasta avea o interpretare care numai cu multă bunăvoinţă putea fi considerată cântat, o ilustraţie muzicală care părea să fi fost desenată cu Microsoft Paint şi un video promoţional care avea motive întemeiate să fie considerat cel mai prost realizat din toate timpurile.5 însă Vanilia avea nişte aliaţi puternici. Datorită câtorva articole publicate în diverse reviste şi al unei apariţii în îndrăgita emisiune *Top of the Pops* a postului BBC, cântecul tot a reuşit să ajungă pe locul 14 în topurile

muzicale.[80]

Este de la sine înţeles că succesul formaţiei a fost de scurtă durată. Când a scos cel de-al doilea single, popularitatea ei deja dispăruse. Un al treilea nu a mai apărut. Toate acestea par să sugereze că dovada socială nu este singurul factor care influenţează gusturile, aşa cum a demonstrat un experiment ulterior al echipei de la Music Lab.

Structura celui de-al doilea studiu a fost, în linii mari, aceeaşi ca a celui dintâi. Însă, de această dată, pentru a testa în ce măsură percepţia referitoare la popularitate devenea o profeţie care se auto-împlinea, cercetătorii au adăugat o chichiţă. De îndată ce topurile din fiecare „lume" s-au stabilizat, ei au oprit experimentul şi au întors lista preferinţelor muzicale cu susul în jos. Noii vizitatori ai playerului audio au văzut hiturile la bază piramidei topului, pe când chixurile de pe ultimele locuri, ajunse în vârf, au început să pară *crème de la crème*[81] din acel top muzical.

Aproape imediat numărul total de descărcări a început să scadă. În momentul în care cântecele de pe primele locuri nu au mai fost atrăgătoare, oamenii şi-au pierdut interesul faţă de cântecele de pe întregul website. Declinurile cele mai abrupte s-au înregistrat la descărcările fiasco-urilor ajunse acum în vârful topurilor. În acest timp, piesele bune, care lâncezeau pe ultimele locuri, au avut o soartă mai nefericită decât atunci când erau în vârf, dar tot s-au descurcat mai bine decât cele care se situaseră, anterior, la sfârşitul listei de preferinţe. Dacă oamenii de ştiinţă ar fi lăsat experimentul să continue suficient de mult timp, melodiile cele mai bune dintre cele mai bune şi-ar fi restabilit popularitatea. În concluzie: piaţa nu este blocată într-o anumită stare. Atât şansa, cât şi calitatea au un rol important de jucat.6

80 Este mare păcat că formatul acestei cărţi nu permite includerea unor fragmente muzicale, pentru că îmi doresc foarte mult să ascultaţi cât de îngrozitor de prost este cântecul acesta. Căutaţi-1 pe Google, bine? (n. a.).

81 „Cele mai bune" (fi*. în original) (n. tr.).

Revenind la realitate – acolo unde s-au strâns suficiente informaţii doar dintr-o singură lume pe care să le putem folosi – există o interpretare simplă a descoperirilor din experimentele efectuate în Music Lab: calitatea contează şi ea nu este totuna cu popularitatea. Faptul că melodiile cele mai bune şi-au recâştigat popularitatea dovedeşte că o anumită muzică este, inerent, „mai bună" decât o alta. La o extremă, un cântec senzaţional al unui artist nemaipomenit ar trebui (cel puţin teoretic) să fie menit succesului. Însă problema este că reciproca nu e neapărat valabilă. Numai pentru că un lucru are succes nu înseamnă că este şi de calitate.

Definirea calităţii este o cu totul altă poveste, la care vom ajunge imediat. Însă pentru unii calitatea în sine nu este neapărat importantă. Dacă eşti o companie de înregistrări, un producător de film sau un grup editorial, întrebarea de un milion de dolari este: poţi să observi cu anticipaţie un succes garantat? Poate fi creat un algoritm care să selecteze hiturile?

VÂNĂTOAREA DE HITURI

Investirea banilor în filme este o afacere riscantă. Puţine sunt filmele care aduc bani; cele mai multe abia dacă îşi scot cheltuielile, iar eşecurile nu pot fi evitate în totalitate.7 Este o afacere cu mizã mare: atunci când costurile realizării unui film ajung la zeci sau sute de milioane de dolari, eşecul în prezicerea cererii pentru acel produs poate fi catastrofal de costisitor.

Este o lecţie pe care au învăţat-o pe propria piele cei de la Disney, cu filmul *John Carter*, lansat în 2012. Studioul a aruncat pe apa Sâmbetei 350 de milioane de dolari pentru realizarea sa şi a decis că acesta ar fi trebuit să se alăture unor *blockbuster-e* precum *Toy Story* şi *Finding Nemo*, devenind primul produs din următoarea lor mare franciză. Cum, nu l-aţi văzut? Ei bine, nici eu! Filmul nu a reuşit să capteze imaginaţia publicului şi a sfârşit prin a produce o pierdere de 200 de milioane de dolari, ceea ce a dus la demisia preşedintelui studiourilor Walt Disney.8

Cei mari şi tari de la Hollywood au acceptat întotdeauna ideea că efectiv nu ai cum să prezici succesul comercial al unui film. Aşa ceva ţine de instinct. A-ţi încerca norocul cu filme care ar putea să dea lovitura la încasări face parte din meserie. În 1978, Jack Valenţi, preşedinte şi director general al Motion Picture Association of America, s-a exprimat în felul următor:

Nimeni nu-ţi poate spune cum va fi primit un film pe piaţă. Nu înainte să ruleze într-o sală de cinematograf cu luminile stinse, iar între ecran şi spectatori să zboare scântei.9

Cinci ani mai târziu, în 1983, William Goldman, care a scris romanele care stau la baza filmelor *The Princess Bride* (*File de poveste*) şi *Butch Cassidy* şi *Sundance Kid*, a fost şi mai succint: „Nimeni nu are habar de nimic”. 10

Aşa cum am văzut pe parcursul acestei cărţi, algoritmii moderni sunt, de obicei, capabili să prezică până şi ceea ce pare imposibil de prezis. De ce ar face filmele excepţie? Poţi să calculezi succesul unui film cu ajutorul venitului obţinut din încasări şi al aprecierilor critice pe care le obţine. Poţi să calculezi tot felul de factori despre structura şi caracteristicile unui film: distribuţia, genul, durata, caracteristicile intrigii şi aşa mai departe. De ce nu s-ar aplica, aşadar, aceleaşi tehnici în încercarea de a găsi capodoperele? Sau pentru a descoperi ce filme sunt menite să aibă un succes uriaş şi încasări pe măsură?

Aceasta a fost ambiţia câtorva studii ştiinţifice recente, care şi-au propus să se inspire din extinsele şi bogatele baze de date strânse şi organizate de site-uri precum Internet Movie Database (IMDB) sau Rotten Tomatoes. Şi – lucru deloc surprinzător! – în aceste informaţii se ascund nenumărate revelaţii.

Să luăm, spre exemplu, studiul condus de Sameet Sreenivasan în anul 2013.11 El şi-a dat seama că, deoarece le ceruse utilizatorilor săi să eticheteze filmele cu ajutorul unor cuvinte-cheie care descriau intriga, IMDB crease un catalog de descriptori uimitor de detaliat care ar fi putut

să arate cum se modificaseră preferinţele publicului în materie de filme de-a lungul timpului. În momentul în care Sameet şi-a început studiul, IMDB catalogase peste 2.000.000 de filme care acopereau un interval de timp de peste 100 de ani, fiecare având mai multe etichete care descriau intriga. Unele cuvintecheie erau descrieri mai generale ale filmului, precum „crimăorganizată” sau» relaţia tată-fiu”. Altele se bazau pe localizări, precum „Manhattan, New York City” sau pe anumite aspecte al intrigii, cum ar fi: „ameninţat cu arma” sau „legat de scaun”.

De sine stătătoare, cuvintele-cheie arătau că interesul nostru faţă de anumite elemente ale intrigii tinde să se manifeste în valuri de popularitate (gândiţi-vă, spre exemplu, la filmele despre al Doilea Război Mondial sau la acelea care abordează subiectul avortului). Va exista un val de lansări în succesiune rapidă pe teme similare, după care va urma o perioadă scurtă de acalmie. Luând în considerare toate aceste etichete, Sreenivasan a reuşit să obţină un punctaj pentru caracterul de noutate al fiecărui film în momentul lansării – un număr cu zecimale, cuprins între zero şi unu – care putea fi comparat cu succesul încasărilor sale.

Dacă un anumit aspect al intrigii sau o caracteristică principală (cum ar fi prezenţa scenelor care prezentau nuduri feminine sau aspecte legate de crima organizată) era un aspect familiar al filmelor mai vechi, cuvântul-cheie avea să-i aducă filmului un punctaj scăzut pentru noutate. Însă caracteristicile originale ale intrigii – cum ar fi introducerea artelor marţiale în filmele de acţiune din anii 1970, să spunem – îi puteau aduce filmului un punctaj ridicat pentru noutate atunci când caracteristica respectivă apărea pentru prima oară pe marile ecrane.

După cât se pare, avem o relaţie complicată cu acest caracter de noutate. În medie, cu cât punctajul pentru noutate cil unui film era mai ridicat, cu atât acesta avea încasări mai mari. Dar numai până la un punct. Dacă

depăşeai pragul de noutate, înaintea ta se deschidea o prăpastie: veniturile obţinute de un film înregistrau o prăbuşire bruscă la orice punctaj care depăşea valoarea de 0,8. Studiul lui Sreenivasan dezvăluia ceva ce sociologii bănuiau demult: ne îndepărtăm de ceea ce este banal, dar, totodată, detestăm ceea ce este complet necunoscut. Filmele cele mai bune se situează într-o nişă îngustă şi confortabilă, între „nou” şi „nu foarte nou”.

Punctajul pentru caracterul de noutate poate fi un mod util de a ajuta studiourile de film să evite susţinerea unor chixuri, dar nu este de prea mare folos dacă doreşti să afli ce soartă va avea un anumit film. Pentru acest lucru ar putea fi mai utilă activitatea unei echipe de cercetători europeni care a descoperit existenţa unei legături între numărul de editări făcute unui film pe o pagină de Wikipedia în luna dinaintea lansării în cinematografe şi posibilele încasări.12 Editările erau, adesea, făcute de persoane fără legătură cu lansarea filmului, de către simpli pasionaţi de film care contribuiau cu informaţii la pagina respectivă. Editările mai numeroase însemnau că în jurul unei lansări se stârnise mai multă vâlvă ceea ce ducea la încasări mai mari.

În ansamblu, modelul lor avea o putere predictivă modestă: din 312 filme studiate, au prezis corect încasările pentru 70 dintre acestea, cu o precizie egală sau mai mare de 70%. Dar cu cât un film era mai bine primit de public, cu atât se făceau mai multe editări pe pagina din Wikipedia şi cu atât echipa de cercetători avea mai multe informaţii pentru a-şi continua studiul, iar, prin urmare, cu atât deveneau mai exacte predicţiile pe care le făcea aceasta. Încasările obţinute de şase filme cu câştiguri mari au fost prezise corect, cu o acurateţe de 99%!

Studiile acestea sunt interesante sub aspect intelectual, însă un model care funcţionează numai cu o lună înainte de lansarea unui film nu este de prea mare folos pentru investitori. Dar ce ar fi să abordăm întrebarea direct: să luăm toţi factorii implicaţi dinainte cunoscuţi –

genul filmului, nivelul de celebritate al actorilor principali, clasificarea pe categorii de vârstă (AP, 12 etc.) – şi să folosim un algoritm de învăţare automată pentru a prezice dacă un film va fi considerat excepţional?

Un studiu celebru din 2005 a făcut exact acest lucru, utilizând o reţea neurală în încercarea de a prezice performanţa unor filme cu mult înainte de lansarea lor în cinematografe.13 Pentru a face lucrurile cât mai simple, autorii săi au renunţat la prezicerea exactă a încasărilor, încercând în schimb să clasifice filmele într-una din nouă categorii, de la eşec răsunător la film cu încasări record. Din păcate, chiar şi cu această etapă simplificatoare, rezultatele obţinute au lăsat mult de dorit. Reţeaua neurală a depăşit performanţele tuturor modelelor statistice folosite anterior, dar tot a reuşit să clasifice corect performanţa unui film, în medie, în numai 36, 9% din cazuri. În categoria de top – filme ale căror încasări au depăşit 200 de milioane de dolari – algoritmul s-a descurcat ceva mai bine, identificând corect filmele de succes în 47, 3% din cazuri. Dar investitori, păziţi-vă! în jur de 10% dintre filmele pe care algoritmul le-a considerat a fi capabile să devină filme de succes au ajuns să aibă încasări sub 20 de milioane de dolari, ceea ce, după standardele hollywoodiene, este o sumă de bani demnă de toată mila.

Celelalte studii întreprinse de atunci au încercat să îmbunătăţească aceste predicţii, însă până în prezent nimeni nu a făcut vreun progres semnificativ. Toate dovezile trimit la o singură concluzie: până când nu ai informaţii despre reacţia celor dintâi spectatori, popularitatea viitoare a unui film este, în mare măsură, imprevizibilă. Când vine vorba despre alegerea filmelor de succes dintr-o multitudine, Goldman avea dreptate: nimeni nu are habar de nimic.

CUANTIFICAREA CALITĂŢII

Prezicerea popularităţii este, aşadar, o chestiune complicată. Nu există o modalitate simplă de a separa ceea

ce ne place de motivul pentru care ne place lucrul respectiv, iar în domeniul creativităţii asta este, mai degrabă, o problemă numai bună pentru algoritmi. Dacă nu poţi folosi popularitatea pentru a-ţi semnala ce este „bun”, cum poţi să măsori atunci calitatea?

Iată un lucru important: dacă vrem ca algoritmii să aibă autonomie în cadrul artelor – fie să creeze opere noi, fie să ne ofere informaţii semnificative despre arta pe care o creăm noi înşine – vom avea nevoie de o modalitate de măsurare a calităţii pentru a merge mai departe. Trebuie să existe un mod obiectiv de a îndruma algoritmul în direcţia corectă, un fel de adevăr întemeiat, empiric indubitabil pe care acesta să-l aibă ca referinţă. Un fel de analogie valabilă în domeniul artei pentru: „acest grup de celule este purtător de cancer” sau „acuzatul a continuat să comită infracţiuni, a recidivat, după eliberarea pe cauţiune”. În lipsa acesteia, progresul este dificil. Nu putem să proiectăm un algoritm sau să găsim un cântec „bun”, dacă nu putem să definim ce înţelegem prin „bun”.

Din nefericire, în încercarea de a găsi o unitate de măsură obiectivă pentru calitate, ne confruntăm cu o întrebare filosofică foarte controversată care datează încă de pe vremea lui Platon, o chestiune care se dezbate de mai bine de 2000 de ani: cum poate fi apreciată valoarea estetică a artei?

Unii filosofi – precum Gottfried Leibniz – susţin că, dacă există obiecte pe care le considerăm cu toţii frumoase (spre exemplu, *David* al lui Michelangelo sau *Lacrimosa* lui Mozart), atunci ar trebui să existe o esenţă definibilă şi măsurabilă a frumuseţii care face ca un obiect de artă să fie, în mod obiectiv, mai bun decât altul.

Pe de altă parte, se întâmplă destul de rar ca toată lumea să fie de aceeaşi părere. Alţi filosofi, cum ar fi David Hume, susţin că „frumuseţea se află în ochiul privitorului”, cum spunem noi, englezii, adică: „nu-i frumos ce e frumos, ci-i frumos ce-mi place mie”. Gândiţi-vă la opera lui Andy Warhol, de exemplu, care unora le oferă o experienţă

estetică intensă, pe când altora li se pare imposibil de distins, sub aspect artistic, de o cutie cu supă.

Alţii, printre care se numără şi Immanuel Kant, au spus că adevărul este pe undeva, pe la mijloc. Că aprecierea frumuseţii nu este în totalitate subiectivă, dar nu poate fi nici cu totul obiectivă. Aprecierea este senzorială, emoţională şi intelectuală totodată – şi, foarte important, ea se poate modifica în timp, în funcţie de starea de spirit a observatorului.

Există, cu siguranţă, dovezi care vin în sprijinul acestei idei. Fanii lui Banksy îşi amintesc, poate, cum a aranjat acesta un stand în Central Park din New York în anul 2013, vânzând în anonimitate picturi originale pe pânză, vopsite cu spray în alb şi negru, pe care a cerut 60 de dolari bucata. Standul era ascuns într-un şir de altele asemănătoare în care turiştilor li se vindeau mărunţişurile obişnuite, aşa că preţul cerut trebuie să li se fi părut trecătorilor destul de mare. Se scurseseră câteva ore până când cineva a decis să cumpere un tablou. În ziua aceea Banksy a câştigat, în total, suma de 420 de dolari.14 Un an mai târziu, la o casă de licitaţii din Londra, un cumpărător a apreciat că valoarea estetică a aceloraşi opere de artă era suficient de mare încât să fie tentat să cheltuiască suma de 68.000 de lire sterline (115.000 de dolari americani la cursul de atunci) pe o singură pictură.15

Este adevărat că nu oricui îi place Banksy (Charlie Brooker – cel care a creat serialul *Black Mirror*[82] – l-a descris odată ca pe „un aiurit [a cărui] operă pare uimitor de deşteaptă pentru proşti”).16 Aşa că s-ar putea susţine că istorioara aceasta este doar dovada faptului că opera lui Banksy nu are o calitate inerentă. Nu este decât o şmecherie la modă (şi o dovadă a mentalităţii de turmă), care atrage nişte preţuri atât de mari, că-ţi vine să plângi. Dar aprecierea noastră estetică, atât de capricioasă, a fost urmărită şi în raport cu forme de artă care sunt de o incontestabilă calitate.

82 Vezi nota de la p. 70 (n. red.).

Exemplul meu preferat este un experiment realizat de *Washington Post* în 2007.17 Ziarul i-a cerut violonistului de faimă internaţională Joshua Bell să mai adauge o reprezentaţie pe lista sa de concerte cu casa închisă. Înarmat cu vioara sa Stradivarius de 3, 5 milioane de dolari, Bell s-a postat lângă o scară rulantă dintr-o staţie de metrou din Washington, DC, într-o dimineaţă, în timpul orei de vârf, a aşezat o pălărie pe jos, să strângă banii donaţi de trecători şi a cântat timp de 43 de minute. Aşa cum a scris *Washington Post*, acolo se afla unul dintre „cei mai buni interpreţi de muzică clasică din lume, care cânta muzica cea mai distinsă care fusese compusă vreodată la una dintre cele mai scumpe viori create vreodată”. Rezultatul? Şapte persoane s-au oprit să-l asculte câteva minute. Mai bine de 1 (XX) au trecut, pur şi simplu, pe lângă el. La sfârşitul reprezentaţiei, Bell strânsese în pălărie suma infimă de 32, 17 dolari.

Ceea ce considerăm noi a fi „bun”, se modifică de asemenea. Apetitul pentru anumite genuri de muzică clasică a fost deosebit de rezistent la trecerea timpului, însă nu se poate spune acelaşi lucru şi despre alte forme de artă. Armand Leroi, profesor de biologie evoluţionistă la Imperial Collège din Londra, a studiat evoluţia muzicii pop şi a descoperit, în cursul analizei, dovezi clare ale modificării preferinţelor noastre faţă de aceasta:

Populaţia posedă intrinsec un prag de plictiseală. Există o tensiune care se acumulează atunci când oamenii simt nevoia de ceva nou.18

Ca să ofer un exemplu, gândiţi-vă la tobele electronice şi sintetizatoarele care au fost la modă în muzica pop a anilor 1980 – atât de la modă, încât diversitatea muzicală din topuri mai-mai c-a dispărut. „Totul suna à la Madonna la începutul carierei sau ca un cântec compus de cei de la Duran Duran”, ne lămureşte Leroi. „Dar poate că vă spuneţi: «Bine! Am ajuns în punctul culminant al muzicii pop. El este chiar acolo. A fost găsită formula supremă a succesului!»” Numai că,

bineînţeles, acest punct nu a fost găsit. La scurt timp, odată cu apariţia hip-hopului, diversitatea muzicală din topurile muzicale a explodat din nou. Avea hip-hopul ceva special, ceva care dusese la o schimbare? ham întrebat eu pe Leroi. Acesta mi-a răspuns: „Nu cred. Ar fi putut fi orice altceva, dar s-a întâmplat să fie genul hip hop. La care consumatorul american a reacţionat, spunând: «Ei bine, iată ceva nou. Daţi-ne mai mult din ăsta.»”.

Ideea este următoarea: chiar dacă există unele criterii obiective care fac ca o operă de artă să fie mai bună decât o alta, atât timp cât contextul joacă, pentru noi, un rol în aprecierea estetică a artei, nu este posibil să creezi o unitate de măsură concretă pentru calitatea estetică, una care să fie valabilă în toate locurile şi în toate timpurile. Indiferent de tehnicile statistice, de scamatoriile inteligenţei artificiale sau de algoritmii de învăţare automată pe care îi întrebuinţezi, încercarea de a folosi cifre pentru a cuprinde esenţa excelenţei artistice este asemenea încercării de a prinde fumul cu mâinile.

Un algoritm are, însă, nevoie de *ceva* pentru a merge mai departe. Aşadar, în momentul în care renunţi la popularitate şi la calitatea inerentă, îţi mai rămâne un singur lucru care să poată fi cuantificat: un indicator de similaritate cu orice a existat mai înainte.

Se pot face încă multe lucruri utilizând indicatorii de similaritate. Când vine vorba de construirea unui motor de recomandări, cum sunt cele de la Netflix şi Spotify, similaritatea este, fără îndoială, indicatorul ideal. Ambele companii au o anumită modalitate de a ajuta utilizatorii să descopere filme şi cântece noi, iar, dacă se abonează la serviciile lor, ambele susţin că vor prezice cu acurateţe preferinţele utilizatorilor. Nu-şi pot seta algoritmii pe ceea ce este la modă în acel moment, pentru că utilizatorii vor fi asediaţi de sugestii de genul Justin Bieber şi *Peppa Pig – The Movie* („Filmul cu Purceluşa Peppa”). Nu şi-i pot seta nici pe vreun fel de substitut pentru calitate, cum ar fi recenziile, pentru că atunci paginile lor ar fi inundate de

manifestări artistice la care adormi de-a-n picioarelea de plictiseală, atunci când oamenii nu-şi doresc decât să-şi scoată încălţările după o zi lungă de muncă şi să se piardă într-un thriller de doi bani sau să se uite la ceva cu Ryan Gosling vreo două ore.

Similaritatea, în schimb, îi permite algoritmului să se concentreze strict pe preferinţele unei persoane. Ce muzică ascultă, ce filme urmăreşte, la ce revine în mod repetat? Apoi, programul poate folosi IMDB, Wikipedia, bloguri de muzică sau articole din reviste pentru a extrage câteva cuvintecheie pentru fiecare cântec, artist sau film. Dacă procedează în acelaşi fel pentru întregul catalog, îi va fi uşor apoi să găsească şi să recomande alte cântece şi filme care au faguri, „etichete descriptive”, asemănătoare. În plus, poate găsi şi alţi utilizatori cărora le-au plăcut filme sau cântece similare, să vadă ce alte cântece şi filme le-au plăcut acestora şi să le recomande primului utilizator.

În niciun moment Spotify sau Netflix nu încearcă să-ţi ofere cântecul sau filmul perfecte. Perfecţiunea le interesează prea puţin. Spotify Discover nu promite că va căuta singura formaţie din lume menită să corespundă în întregime şi negreşit cu gusturile şi cu dispoziţia ta actuală. Recomandările algoritmilor îţi oferă doar cântecele şi filmele care sunt suficient de bune încât să te împiedice să fii dezamăgit. Îţi oferă o modalitate inofensivă de a-ţi petrece timpul liber. Din când în când, vor găsi ceva care îţi va plăcea la nebunie, dar asta este un fel de scamatorie din partea lor, o păcăleală similară aceleia folosite de şarlatanii care practică ghicitul în palmă şi care pot descoperi multe lucruri despre o persoană nu din liniile desenate în palma acesteia, ci observându-i comportamentul, vestimentaţia etc. Nu este necesară decât o reuşită, din când în când, pentru a crea impresia că ai avut norocul să descoperi o muzică nouă. Motoarele de căutare nu simt nevoia să aibă dreptate de fiecare dată.

Similaritatea funcţionează de minune pentru motoarele de căutare. Când le ceri, însă, algoritmilor să

creeze artă în lipsa unui etalon clar pentru calitate, ei bine, abia atunci lucrurile încep să devină interesante. Poate un algoritm să fie creativ dacă singura sa noţiune despre artă ţine de ceva ce s-a petrecut în trecut?

„ARTIŞTII BUNI COPIAZĂ, ARTIŞTII MARI FURĂ” – PABLO PICASSO

În octombrie 1997, la Universitatea din Oregon au venit nişte spectatori pentru a asista la un concert neobişnuit. Pe scenă, în faţa lor, era aşezat un singur pian. Apoi a luat loc în faţa sa pianista Winifred Kemer, care a început să cânte trei bucăţi muzicale scurte, distincte.

Una a fost o compoziţie pentru pian mai puţin cunoscută scrisă de maestrul barocului, Johann Sebastián Bach. A doua a fost compusă în stilul lui Bach de Steve Larson, profesor de muzică la acea universitate. Iar cea de-a treia a fost compusă de un algoritm conceput în mod intenţionat astfel încât să imite stilul lui Bach.

După ce au ascultat toate cele trei piese, spectatorilor li s-a cerut să ghicească în mod exact cui îi aparţinea fiecare piesă. Spre stupoarea lui Steve Larson, majoritatea a considerat că piesa lui era cea care fusese, în realitate, compusă de calculator. Iar în oftaturile unanime de consternare încântată, spectatorilor li s-a dezvăluit că muzica pe care o consideraseră ca aparţinându-i lui Bach nu era decât creaţia unei maşini.

Larson nu a fost deloc mulţumit. Într-un interviu acordat ziarului *New York Times* la scurt timp după experiment, el a declarat:

Admiraţia mea faţă de muzica lui Bach este profundă, cosmică. Am fost foarte tulburat când mi-am dat seama că oamenii puteau fi păcăliţi de un program de calculator.

Nu era singurul surprins neplăcut. David Cope, bărbatul care concepuse remarcabilul algoritm care a stat la baza bucăţii muzicale compuse de calculator, a mai întâlnit această reacţie, căci, după cum mi-a mărturisit:

[La început] am jucat ceea ce eu numeam „jocul” cu persoanele”. Iar când acestea greşeau, se înfuriau. Erau,

oricum, supărate pe mine pentru că inventasem tot acel concept. Deoarece creativitatea este considerată o însuşire specific umană.19

Aceasta fusese şi opinia lui Douglas Hofstadter, profesor de ştiinţe cognitive şi autorul a numeroase cărţi, care organizase concertul de la bun început. Cu câţiva ani înainte, în 1979, când i se decernase Premiul Pulitzer pentru cartea sa *Godel Escher, Bach**, Hofstadter îşi exprimase ferm opinia în această privinţă:

Muzica este un limbaj al emoţiilor şi până când programele [de calculator] nu vor avea emoţii tot atât de complexe ca ale noastre, este imposibil ca un program să compună ceva frumos… Să te gândeşti că am putea să-i cerem unui „tonomat" dinainte programat să născocească piese pe care le-ar fi putut scrie Bach este o subestimare grotescă şi ruşinoasă a profunzimilor spiritului uman.20

După ce a ascultat, însă, produsul algoritmului lui Cope, EMI (acronimul pentru *Experiments în Musical Intelligence*, „Experimente asupra Inteligenţei Muzicale"), Hofstadter a recunoscut că, poate, lucrurile nu erau chiar atât de simple. Iată ce a mărturisit el în zilele care au urmat după experimentul de la Universitatea din Oregon:

— Publicată în limba română în traducerea Corinei Tiron, cu titlul *Godel, Escher, Bach. Brilianta ghirlandă eternă*, Editura Humanitas, Bucureşti, 2015 (n.tr.).

Sunt uimit şi tulburat de creaţiile lui EMI. Singura mea consolare în acest moment vine din faptul că mi-am dat seama că EMI nu generează stil de la sine, căci este dependent de imitarea compozitorilor precedenţi. Dar nici asta nu este o consolare prea mare. Spre absoluta mea deznădejde, [poate că] muzica înseamnă mult mai puţin decât am crezut eu vreodată.21

Prin urmare, care este adevărul? Excelenţa estetică este apanajul activităţii umane? Sau un algoritm este capabil să creeze artă? Şi dacă nişte spectatori nu au reuşit să deosebească muzica lui EMI de aceea a unui mare maestru înseamnă că maşina a dat dovadă de o

autentică creativitate?

Să încercăm să abordăm, pe rând, aceste întrebări, începând cu ultima. Pentru formarea unei opinii pertinente şi bine documentate, merită să ne oprim puţin pentru a înţelege cum funcţionează algoritmul[83] David Cope a fost destul de generos, încât să mă lămurească în această privinţă.

Primul pas în construirea unui algoritm a fost traducerea muzicii lui Bach în ceva care să poată fi înţeles de un calculator:

Trebuie să introduci într-o bază de date cinci reprezentări ale unei singure note: timpul de aplicare, durata, înălţimea sunetului, volumul sonor şi instrumentul.

Pentru fiecare notă din compoziţiile lui Bach, Cope a trebuit să introducă în calculator, meticulos, de mână, aceste cinci reprezentări numerice şi gândiţi-vă că Bach a compus 371 de corale! Dar există multe alte piese, care însumează zeci de mii de note şi pentru fiecare dintre acestea au fost introduse cele cinci cifre! A fost nevoie de un efort monumental din partea lui Cope:

Luni întregi nu am făcut altceva, zi de zi, decât să introduc numere în calculator. Dar trebuie să mărturisesc: eu pot să devin complet obsedat de ceva.

Apoi Cope a luat fiecare timp, fiecare bătaie, din muzica lui Bach şi a studiat ce se întâmpla ulterior. După fiecare notă interpretată într-o corală de Bach, Cope a înregistrat nota imediat următoare. A aşezat totul laolaltă într-un fel de dicţionar – o bancă de date în care algoritmul putea să caute un singur acord şi să găsească o listă exhaustivă a tuturor direcţiilor în care condeiul lui Bach trimisese muzica în continuare.

Judecând din acest punct de vedere, EMI seamănă oarecum cu algoritmii pentru generarea textului predictiv

83 Aş dori, de asemenea, să vă recomand insistent să căutaţi măcar o parte din muzica lui Cope care există on-line. Cred că preferata mea este piesa orchestrală în stilul lui Vivaldi, care a fost postată pe YouTube la 12 august 2012: https://www.youtube.com/watch?v=2kuY3BrmTfQ (link accesat cel mai recent la 28 martie 2019).

din *smartphone*-ul dumneavoastră. Bazându-se pe enunţurile pe care le-aţi scris în trecut, telefonul creează un dicţionar de cuvinte pe care este posibil să doriţi să le scrieţi în continuare şi vi le oferă ca sugestie pe măsură ce vă compuneţi textul.[84] Etapa finală a fost să lase maşina de capul ei. Cope alimenta sistemul cu un acord de început şi instruia algoritmul să-l caute în dicţionarul său şi să decidă ce să cânte după aceea, alegând un nou acord la întâmplare din acea listă. Apoi algoritmul repeta procedura – căutând fiecare acord ulterior în dicţionar pentru a alege următoarele note pe care să le cânte.

Rezultatul l-a constituit o compoziţie cu totul originală care sună ca şi când ar fi fost compusă de Bach însuşi.[85]

Sau poate chiar *este* Bach. În orice caz, aceasta este viziunea lui Cope:

Bach a compus toate acordurile. Este ca şi când ai lua o bucată de brânză parmezan şi ai da-o pe râzătoare, după care ai încerca să o reconstitui. Tot parmezan va rămâne.

Indiferent cui i se cuvine meritul final, un lucru este sigur. Oricât de frumos ar suna muzica creată de EMI, aceasta se bazează pe simpla recombinare a operei existente. Ea imită tiparele existente în muzica lui Bach, în loc să compună, efectiv, o piesă muzicală diferită.

Mai recent, au fost creaţi şi alţi algoritmi care produc o muzică plăcută sub aspect estetic, ceea ce este un pas înainte faţă de simpla recombinare. O abordare de mare

84 Algoritmul de text predictiv poate fi, de asemenea, folosit pentru a „compune” un SMS în stilul dumneavoastră. Nu trebuie decât să deschizi o pagină nouă şi să îi oferi algoritmului câteva cuvinte de pornire, cum ar fi „m-am născut”. După aceea selectezi, în mod repetat, cuvintele care apar pe ecran. Iată ce mi-a ieşit mie (sincer!). începe destul de bine, dar devine puţin forţat pe parcurs: „(m-am născut) ca să fiu un om bun şi aş fi fericită să fiu cu tine o mulţime de persoane ştiu că nu-mi primeşti e-mail-urile şi nu am deloc timp pentru aşa ceva” (n. a.)

85e xemplu, s-au dovedit a corespunde cu o compoziţie pe care ar fi compus-o însuşi Bach (n. a.).

succes a reprezentat-o folosirea algoritmilor genetici – un alt tip de învăţare automată, care încearcă să exploateze modul în care funcţionează selecţia naturală. În definitiv, dacă privim păunii înseamnă că evoluţia ştie câte ceva despre crearea frumuseţii.

Ideea este simplă: în interiorul acestor algoritmi, notele sunt prelucrate ca şi când ar fi un fel de ADN al muzicii. Totul începe de la o populaţie iniţială de „cântece” – fiecare fiind un amalgam de note, luate la întâmplare şi puse laolaltă. După multe „generaţii”, algoritmul se inspiră din cântece, găsind şi răsplătind elementele „frumoase” din muzică pentru a da naştere, cu timpul, unor compoziţii din ce în ce „mai bune”. Spun „frumoase” şi „mai bune” însă, aşa cum deja ştim, nu există nicio modalitate de a stabili ce înseamnă exact oricare din aceste cuvinte. Algoritmul poate să creeze poeme şi picturi tot atât de bine ca şi muzică, numai că nu dispune – încă! – pentru a continua, decât de un anumit grad de similitudine cu tot ce a fost creat mai înainte.

Iar uneori asta este tot ce-ţi trebuie. Dacă eşti în căutarea unei melodii de fundal pentru site-ul tău sau pentru un video pe YouTube care să sune, în linii mari, ca o piesă folk, nu-ţi pasă că ea seamănă cu cele mai bune cântece folk din trecut. De fapt, nu-ţi doreşti decât ceva care să nu încalce drepturile de autor, şi să nu-ţi dea bătaia de cap de a trebui să compui ceva tu însuţi. Iar dacă asta este tot ce doreşti să obţii, există câteva companii care te pot ajuta. Firmele britanice Jukebox şi Al Music oferă deja acest gen de servicii, utilizând algoritmi care sunt capabili să creeze muzică. O parte din aceasta se va dovedi utilă. O parte va fi (aproape) originală. O parte va fi chiar frumoasă. Algoritmii sunt, fără îndoială, imitatori excelenţi, dar nu şi inovatori foarte buni.

Nu afirm acest lucru pentru a le face un deserviciu algoritmilor. Cea mai mare parte a muzicii compuse de fiinţele umane nu este, nici ea, deosebit de inovatoare. Dacă ar fi să-l întrebăm pe Armand Leroi, biologul

evoluţionist care a studiat evoluţia culturală a muzicii pop, suntem puţin cam idealişti în privinţa inventivităţii umane. Până şi succesele deosebite din topurile muzicale, spune el, ar putea fi generate de o maşină. Iată, de exemplu, ce părere are despre cântecul „Happy" al lui Pharrell Williams (ceva îmi spune că nu este un fan al acestui artist!):

Mappy, happy, happy, I'm so happy!" * Zău aşa! Are, să spunem, cam şase cuvinte în textul ăsta. Este un cântec cum nu se poate mai robotic, care descurajează chiar şi dorinţa omenească cea mai înălţătoare şi fundamentală de a asculta o melodie optimistă de vară. Este cântecul cel mai idiot şi simplist care poate fi încă acceptat. Iar dacă *acesta* este nivelul – ei bine, atunci nu este prea dificil pentru un algoritm să-l creeze.

Leroi nu are o părere prea bună nici despre calităţile lirice ale lui Adele:

Dacă ar fi să analizezi oricare dintre cântecele ei, nu ai găsi niciun sentiment care să nu poată fi creat de un generator de cântece triste.

* „Fericit, fericit, fericit, / Sunt atât de fericit!" (engl.) (n.tr.).

S-ar putea să nu fiţi de aceeaşi părere (nici eu nu sunt sigură că sunt), dar este, fără îndoială, un argument în favoarea faptului că totuşi cea mai mare parte a creativităţii umane – cum sunt şi produsele algoritmilor care „compun" – nu este decât o recombinare a unor idei preexistente. Aşa cum spunea Mark Twain:

Nu există ceva care să se numească „o idee nouă". Noi nu facem decât să luăm, pur şi simplu, o mulţime de idei vechi şi le aşezăm într-un fel de caleidoscop mental. Le învârtim puţin, iar ele alcătuiesc combinaţii noi şi interesante. Le învârtim în continuare şi obţinem combinaţii noi la nesfârşit; numai că ele sunt aceleaşi vechi fragmente de sticlă colorată care au fost utilizate de-a lungul veacurilor.22

Pe de altă parte, Cope are o definiţie a creativităţii

foarte simplă care înglobează cu uşurinţă ceea ce pot algoritmii să facă:

Creativitatea înseamnă doar să găseşti o asociere între două lucruri care, în mod obişnuit, nu par să aibă ceva în comun.

Poate că aşa este. Dar eu, una, tot simt că, dacă EMI şi algoritmii asemenea lui chiar *dau* dovadă de creativitate, atunci este o formă destul de firavă de creativitate. Muzica lor poate fi frumoasă, dar nu este profundă. Şi oricât m-aş strădui, nu prea pot să scap de sentimentul că, dacă privim produsele acestor maşini drept artă, atunci rămânem cu o viziune despre lume destul de sărăcită cultural. Sub aspect cultural poate fi o „mâncare reconfortantă". Însă nu va deveni niciodată artă cu A mare.

Pe când mă documentam pentru scrierea acestui capitol, mi-am dat seama că motivul nemulţumirii pe care o resimţeam faţă de algoritmii care produc artă îşi are originea într-o întrebare diferită: adevărata problemă nu este dacă maşinăriile pot fi creative (pentru că pot fi), adevărata problemă este ce se înţelege, de fapt, prin artă.

Eu sunt matematiciană. Pot, cu toată încrederea, să fac schimb de informaţii despre rezultate fals-pozitive şi adevăruri absolute referitoare la gradele de acurateţe şi la statistici. În domeniul artistic, însă, aş prefera să ţin cont de părerea lui Lev Tolstoi. Consider, ca şi el, că adevărata artă are în centrul ei legăturile umane şi comunicarea emoţiilor. Aşa cum a spus Tolstoi:

Arta nu este o abilitate meşteşugărească, ea înseamnă transmiterea sentimentelor resimţite de artist.23

Dacă sunteţi de acord cu argumentul adus de el, atunci există un motiv pentru care calculatoarele nu vor putea niciodată să producă artă autentică. Un motiv foarte frumos a fost exprimat de Douglas Hofstadter cu ani înainte de a se fi întâlnit cu EMI:

Un „program" care să producă muzică [...] ar trebui să colinde de unul singur în jurul lumii, croindu-şi drum prin labirintul vieţii şi trăind fiecare clipă. Ar trebui să

înţeleagă bucuria şi singurătatea transmise de vânt într-o noapte friguroasă, să tânjească după mâna iubitei, să înţeleagă inaccesibilitatea unui oraş îndepărtat, frângerea inimii şi revenirea ei la viaţă după o trecere în nefiinţă. Ar trebui să fi cunoscut resemnarea şi dezgustul de viaţă, suferinţa şi disperarea, hotărârea şi reuşita, pietatea şi venerarea. Ar trebui să fi amestecat sentimente opuse, precum nădejdea şi teama, neliniştea şi jubilarea, seninătatea şi aşteptarea.

Din el ar trebui să facă parte un simţ al graţiei, umorului, ritmului, o senzaţie de neprevăzut – şi, bineînţeles, o minunată conştientizare a magiei unei proaspete creaţii, în acestea şi numai în acestea se adăpostesc izvoarele înţelegerii în muzică.24

S-ar putea, prea bine, să greşesc. Poate că, dacă arta algoritmilor începe să semene a creaţie umană autentică (aşa cum s-a întâmplat cu EMI), o vom aprecia totuşi şi îi vom adăuga propriul nostru înţeles. În definitiv, trecutul îndelungat al muzicii pop contrafăcute pare să indice că oamenii pot avea o reacţie emoţională la ceva care nu face decât să semene cu o legătură autentică. Şi, poate, în momentul în care aceste opere de artă algoritmice vor deveni un loc comun, iar noi ne vom da seama că arta nu emană de la o fiinţă umană, această legătură unidirecţională nu ne va mai deranja. La urma-urmei, oamenii leagă relaţii pline de simţire cu obiecte care nu le întorc afecţiunea – mult-iubiţii urşi de pluş din copilărie sau păianjenii transformaţi în animale de companie, de exemplu.

Însă pentru mine, arta adevărată nu poate fi creată din întâmplare. Algoritmii nu pot să ajungă chiar oriunde. Nu chiar orice poate fi cuantificat. Printre toate lucrurile absolut impresionante, năucitoare, pe care mi le pot transmite informaţiile şi statisticile, nu se numără şi cum este să fii om.

ÎNCHEIERE

Rahinah Ibrahim era o arhitectă cu patru copii şi un

soţ care trăia în străinătate, muncea ca voluntar la un spital din localitate şi, pe deasupra, care mai şi lucra să-şi termine doctoratul la Universitatea Stanford. Ca şi când viaţa ei nu ar fi fost destul de ocupată, tocmai suferise o histerectomie de urgenţă şi – deşi între timp îşi revenise destul de bine – întâmpina încă probleme când trebuia să stea în picioare, fără sprijin, o perioadă oarecare, în lipsa medicamentelor. Qi toate acestea, când s-a apropiat cea de-a douăzeci şi opta International Conference on Sistem Sciences[86], în ianuarie 2005, şi-a rezervat un bilet de avion spre Hawaii şi s-a pregătit să-şi prezinte lucrarea cea mai recentă colegilor ei din mediul ştiinţific.1

Dimineaţa devreme, pe 2 ianuarie 2005, când Ibrahim a ajuns la aeroportul din San Francisco împreună cu fiica ei, s-a apropiat de ghişeu, şi-a prezentat actele şi a întrebat-o pe angajata aeroportului dacă o putea ajuta să-şi procure un scaun cu rotile. Membrii personalului nu au fost serviabili. Numele ei a apărut pe ecranul calculatorului strălucind intermitent, semnalând că ar figura pe lista federală a celor care au interdicţie de zbor, o bază de date pusă la punct după 11 septembrie 2001, pentru a-i împiedica pe suspecţii de terorism să călătorească.

Fiica adolescentă a lui Ibrahim, rămasă singură şi năucită lângă ghişeu, a sunat un prieten de familie şi i-a spus că mama ei tocmai fusese luată cu forţa de acolo după ce o încătuşaseră, între timp, Ibrahim fusese aşezată pe bancheta din spate a unei maşini de poliţie şi dusă la secţie. Au percheziţionat-o sub *hijab*[87], au refuzat să o lase să-şi ia medicamentele şi au închis-o într-o celulă. Două ore mai târziu, un agent al Homeland Security[88] a sosit cu

86 Conferinţă Internaţională anuală despre Ştiinţa Sistemelor (engl.) (n. tr.).

87 „Văl” (ar. în original), obiect vestimentar tradiţional al femeilor musulmane, voal care se poartă pe cap şi care lasă faţa vizibilă (n. tr.).

88 United States Department of Homeland Security (Departamentul de Securitate Internă al Statelor Unite), agenţie guvernamentală a SUA similară ministerelor de interne din alte ţări (n. red.).

documentele de eliberare şi i-a spus că fusese scoasă de pe lista celor cărora le era interzis să zboare. Ibrahim a ajuns la conferinţa din Hawaii, după care a luat avionul către Malaysia natală, să-şi viziteze familia.

Ibrahim fusese trecută pe lista celor care aveau interdicţie de zbor când un agent FBI bifase căsuţa greşită pe un formular. S-ar putea ca eroarea să se fi datorat unei confuzii între Jemaah Islamiyah, o organizaţie teroristă cunoscută pentru atacurile cu bombă din Bali, din 2002, şi jemaah Islam, o organizaţie profesională malaysiană pentru persoane care studiază în străinătate. Ibrahim era membră a celei din urmă şi nu avusese niciodată vreo legătură cu cea dintâi. Nu era decât o greşeală, însă una care a avut consecinţe dramatice. Imediat ce eroarea fusese introdusă în sistemul automat, ea dobândise o aură de autoritate care o făcuse imună la orice recurs. Întâlnirea din San Francisco nu a însemnat însă şi finalul poveştii.

La întoarcerea din călătorie, două luni mai târziu, când revenea în SUA din Malaysia, Ibrahim a fost din nou oprită la aeroport. De data aceasta, soluţionarea nu a venit atât de rapid. Îi fusese anulată viza de şedere în SUA pe motiv că fusese suspectată de legături cu o organizaţia teroristă. Cu toate că era mama unui cetăţean american, că locuia în

San Francisco şi activa la una dintre universităţile cele mai prestigioase din ţară, lui Ibrahim nu i se permitea să revină în Statele Unite. Urma să fie nevoită să lupte aproape zece ani pentru a obţine câştig de cauză şi a-şi reabilita numele. Aproape zece ani în care i s-a interzis să pună piciorul pe sol american. Şi toate acestea, din cauza unei erori umane şi a unei maşini cu autoritate nelimitată.

OM PLUS MAŞINĂ

Nu se poate pune la îndoială impactul profund pozitiv pe care l-a avut automatizarea în viaţa noastră. Algoritmii pe care i-am construit până în prezent s-ar putea lăuda cu o listă de realizări uimitor de impresionantă. Ei ne pot

ajuta să diagnosticăm cancerul mamar, să prindem ucigaşi în serie şi să evităm prăbuşirea avioanelor; tot ei ne oferă accesul gratuit şi facil la întreaga bogăţie a cunoaşterii omeneşti şi pun în legătură instantaneu oameni de pe tot globul într-o manieră la care strămoşii noştri nu puteau decât să viseze. Numai că, în insistenţa noastră de a automatiza, în graba noastră de a rezolva multe dintre problemele lumii, se pare că am schimbat o problemă cu alta. Algoritmii – oricât de utili şi de impresionanţi ar fi – ne-au lăsat un ghem tare încâlcit de complicaţii de rezolvat.

Oriunde priveşti – în sistemul judiciar, în sistemul de sănătate, în poliţie, chiar şi în domeniul cumpărăturilor Online – există probleme de confidenţialitate, scăpări, erori, probleme legate de responsabilitate şi de transparenţă care nu vor dispărea foarte uşor. Dacă luăm în considerarea doar algoritmii care există în prezent, ne confruntăm cu probleme de echitate care ajung până la esenţa a cine suntem ca fiinţe umane, cum ne dorim să arate societatea în care trăim şi până unde ne putem adapta la autoritatea iminentă a unei tehnologii lipsite de simţire.

Dar poate că tocmai aşa şi trebuie. Poate că greşim atunci când privim algoritmii ca şi cum ar reprezenta cine ştie ce autoritate.

În primul rând, reticenţa noastră de a pune sub semnul întrebării puterea pe care o are un algoritm a deschis uşa oamenilor care doresc să ne exploateze. Documentându-mă pentru această carte, am întâlnit tot felul de vânzători de iluzii dornici să speculeze mituri şi să profite de credulitatea noastră, în pofida convingătoarelor dovezi ştiinţifice care susţin contrariul, există persoane care vând forţelor de poliţie şi guvernelor lumii algoritmi despre care pretind că ar „prezice” dacă cineva este terorist sau pedofil, exclusiv pe baza trăsăturilor faciale ale acelei persoane! Altele insistă că algoritmul lor poate să sugereze modificarea unei singure replici dintr-un scenariu

care va reuşi să ducă, ea singură, la încasări mult mai mari ale filmului. * Alte persoane afirmă cu îndrăzneală – şi fără nicio urmă de sarcasm – că algoritmul lor poate să-ţi găsească singura şi adevărata iubire.**

— M-am întâlnit, de fapt, cu preşedintele acestei companii pentru a-i lua un interviu. L-am întrebat dacă şi-a validat vreodată algoritmii, să vadă dacă aceştia se ridicau la nivelul aşteptărilor, iar el s-a lansat într-o poveste lungă despre modul în care analiza efectuată de reţeaua neurală a făcut ca o vedetă importantă de la Hollywood să fie exclusă dintr-o serie de filme. Când am remarcat că asta însemna doar că oamenii credeau în algoritmul său şi nu că acesta funcţiona, el a răspuns: „Ei bine, doar nu facem un experiment ştiinţific” (n.a.).

** Există o şmecherie la care se poate apela pentru a descoperi algoritmii care nu-s buni de nimic. Mie îmi place să o numesc Testul Magiei. Ori de câte ori auziţi ceva despre un algoritm, vedeţi dacă puteţi scoate oricare dintre sintagmele la modă, cum ar fi „învăţare automată”, „inteligenţă artificială (IA)”, „reţea neurală”, înlocuind-o cu următorul cuvânt: „magie”. Ceea ce rezultă are încă sens din punct de vedere gramatical? Sau ceva şi-a pierdut înţelesul? Dacă nu mai are noimă, eu mi-aş face griji că, pe undeva, ceva miroase a… ştiţi ce. Fiindcă mă tem că nici măcar în viitorul foarte îndepărtat nu vom „eradica foametea din lume cu ajutorul «magiei»” şi nici nu vom

Dar chiar şi algoritmii care fac ceea ce susţin că fac îşi folosesc adesea în mod incorect autoritatea. Cartea este plină de istorioare despre răul pe care-l pot face aceştia: „instrumentul bugetar” folosit pentru a reduce în mod arbitrar ajutorul financiar pentru persoanele cu dizabilităţi din statul Idaho. Algoritmii pentru depistarea infractorilor care aveau să recidiveze şi care, pe baza informaţiilor din trecutul istoric, sunt mai înclinaţi să sugereze un punctaj de risc ridicat pentru inculpaţii de culoare… Sistemul de depistare a afecţiunilor renale care sileşte milioane de oameni să împărtăşească informaţiile cele mai personale şi

mai intime fără consimţământul sau ştirea lor… Algoritmul din supermarket, care îi răpeşte unei adolescente prilejul de a-şi anunţa tatăl că a rămas însărcinată… Lista cu subiecţi strategici, menită să ajute victimele atacurilor armate, dar care a fost folosită de poliţie ca o listă cu suspecţii buni de arestat… Exemple de inechitate se întâlnesc pretutindeni.

Cu toate acestea, sublinierea deficienţelor algoritmilor poate sugera că există o alternativă perfectă spre care să tindem. M-am gândit mult şi m-am străduit să găsesc fie şi un singur exemplu de algoritm absolut echitabil. Chiar şi aceia care arată bine la prima vedere (cum ar fi sistemul autopilot în cazul avioanelor sau reţelele neurale care diagnostichează cancerul) au probleme dacă îi iei la bani mărunţi. Aşa cum aţi citit în capitolul „Automobilele autonome”, sistemul autopilot îi poate face pe cei care s-au antrenat prin intermediul sistemului automat să aibă un mare dezavantaj atunci când sunt la volan sau când manevrează ei înşişi manşa unui avion. Există chiar o anumită îngrijorare că algoritmii de depistare a tumorilor, pe care i-am întâlnit în capitolul „Medicina” şi care par atât de miraculoşi, nu funcţionează chiar impecabil pentru pacienţii din toate grupurile etnice. Însă exemplele de

„Folosi «magia» pentru a scrie scenariul perfect”, aşa cum nu reuşim să facem toate aceste lucruri nici cu ajutorul IA (n.a.).

sisteme perfect echitabile, corecte, nu prea abundă nici când *nu* sunt implicaţi algoritmii. Oriunde te-ai uita, orice domeniu ai studia, dacă sapi suficient de adânc în interiorul oricărui sistem, vei descoperi un fel sau altul de discriminare.

Imaginaţi-vă, aşadar, pentru o clipă: ce-ar fi să acceptăm faptul că perfecţiunea nu există? Algoritmii chiar vor *face greşeli*. Algoritmii chiar vor *fi nedrepţi*. Asta nu ar trebui să ne împiedice în niciun fel să încercăm să-i facem mai precişi şi mai puţin discriminatorii ori de câte ori ne

stă în putinţă, dar, probabil, recunoscând că algoritmii nu sunt perfecţi, aşa cum nu sunt nici oamenii, s-ar diminua orice prezumţie în privinţa autorităţii de care dispun.

Închipuiţi-vă că, în loc să ne concentrăm atenţia exclusiv asupra construirii unor algoritmi care să corespundă unui standard imposibil de corectitudine desăvârşită, i-am realiza astfel încât să le facilităm remedierea atunci când, inevitabil, vor da greş; că am dedica la fel de mult timp şi efort ca să ne asigurăm că sistemele automate vor fi tot atât de uşor de contestat, pe cât sunt de implementat. Poate că soluţia este să construim algoritmi care să fie discutabili de la bun început. Ce ar fi dacă i-am concepe să asiste oamenii în luarea deciziilor, în loc să le dea instrucţiuni? Să fie transparenţi în privinţa modului în care au ajuns la o anumită concluzie, în loc să ne aducă doar la cunoştinţă un anumit rezultat.

În opinia mea, algoritmii cei mai buni sunt cei care ţin cont de fiinţa umană în fiecare stadiu. Cei care ştiu că avem obiceiul de a manifesta prea multă încredere în produsul unei maşini, recunoscându-şi, totodată, propriile defecte şi afişându-şi incertitudinea cu mândrie, punând-o în centrul atenţiei tuturor.

Aceasta era una dintre cele mai bune particularităţi ale calculatorului Watson al IBM, câştigătorul emisiunii-concurs *Jeopardy*. Cu toate că formatul concursului însemna că maşina trebuia să se oprească la un singur răspuns, algoritmul prezenta şi o serie de alternative pe care le luase în considerare pe parcurs, alături de un punctaj care arăta cât de multă încredere avea în corectitudinea fiecărui răspuns. Dacă punctajele de calcularea a probabilităţii ca un infractor să recidiveze ar include ceva similar, poate că judecătorilor le-ar fi mai uşor să pună sub semnul întrebării informaţia oferită de algoritm. Iar dacă algoritmii de recunoaştere facială ar prezenta un număr de potriviri posibile în loc să aleagă o singură faţă, poate că identificarea greşită nu ar mai

reprezenta o problemă atât de mare.

Aceeaşi caracteristică face să fie atât de eficiente reţelele neurale care scanează lamelele cu ţesut mamar canceros. Algoritmul nu decide ce pacienţi au tumori. El reduce marea varietate de celule, la doar câteva zone suspecte pe care să le verifice patologul. Algoritmul nu oboseşte niciodată, iar patologul rareori emite un diagnostic greşit. Algoritmul şi omul lucrează în parteneriat, profitând fiecare de calităţile celuilalt şi acceptându-şi reciproc defectele.

Mai sunt şi alte exemple – inclusiv din lumea şahului, unde am început această carte. De când a pierdut în faţa lui Deep Blue, Garry Kasparov nu a întors spatele calculatoarelor. Chiar dimpotrivă. A devenit, în schimb, un mare susţinător al ideii de „Centaur Chess", în care un jucător uman şi un algoritm colaborează pentru a concura cu o altă echipă hibridă. Algoritmul evaluează consecinţele posibile în cazul fiecărei mutări, reducând şansele jucătorului de a face o gafă, în timp ce fiinţa umană rămâne la conducerea jocului.

Iată cum descrie Kasparov acest lucru:

Când jucăm asistaţi de calculator, ne putem concentra asupra planificării strategiei în loc să pierdem atât de mult timp făcând calcule. Creativitatea umană devine esenţială în aceste condiţii.24

Rezultă de aici un şah jucat la un nivel mai înalt decât oricând. Un joc tactic desăvârşit şi strategii frumoase şi pline de semnificaţie. Tot ce este cel mai bun din ambele lumi.

Acesta este viitorul pe care sper eu că voi ajunge să-l trăiesc. Unul în care algoritmii aroganţi şi dictatoriali, care umplu multe dintre aceste pagini, vor face parte din trecut. Unul în care vom înceta să privim maşinile ca pe nişte stăpâni obiectivi şi vom începe să le tratăm aşa cum am face-o cu oricare altă sursă de putere. Punându-le deciziile sub semnul întrebării, cercetându-le motivele, recunoscându-ne emoţiile, cerând să ştim cine urmează să

obţină un anumit profit, considerându-le responsabile de propriile greşeli şi refuzând să ne complăcem într-o anumită situaţie. Cred că aceasta este soluţia pentru un viitor în care efectul global al algoritmilor să fie o forţă benefică pentru societate. Şi este corect ca această sarcină să apese pe umerii noştri. Pentru că un lucru este sigur: niciodată nu au fost oamenii mai importanţi decât în era maşinilor, a algoritmilor.

NOTE

Introducere

1. Robert A. Care, The Power Broker: Robert Moses and the Fall of New York, Bodley Head, Londra, 2015, p. 318.

2. Există vreo două eseuri nemaipomenite despre acest subiect care merită citite. Mai întâi, Langdon Winner, „Do artifacts have politics?”, *Daedalus*, vol. 109, nr. 1/1980, pp. 121 – 136 (on-line la https://www.jstor.org/stable/20024652, site accesat cel mai recent la 20 februarie 2019), care conţine exemplul cu podul lui Moses. Şi o versiune mai modernă: Kate Crawford, „Can an algorithm be agonistic? Ten scenes from life în calculated publics”, în *Science, Technology and Human Values*, vol. 41, nr. 1/2016, pp. 77 – 92.

3. *Scunthorpe Evening Telegraph*, numărul din 9 aprilie 1996.

4. Chukwuemeka Afigho (nkeise) a postat pe Twitter un scurt videoclip despre acest efect. Merită să-l căutaţi, dacă nu l-aţi văzut încă. Apare şi pe YouTube, la adresa: https://www.youtube.com/watch?v=87QwWpzvy7I (accesată cel mai recent la 20 februarie 2019).

5. Interviu acordat de Mark Zuckerberg postului de televiziune CNN la 21 martie 2018: „I’m really sorry that this happened”, care a fost urcat şi pe *YouTube*, putând fi vizionat la adresa https://www.youtube.com/watch?v=G6DOhiobfyy (accesată cel mai recent la 20 februarie 2019).

CAPITOLUL 1 – *Puterea*

1. Informaţie obţinută în timpul unei conversaţii particulare cu marele maestru de şah Jonathan Rowson.

2. Feng-Hsiung Hsu, „IBM's Deep Blue Chess grandmaster chips", *IEEE Micro*, vol. 19, nr. 2/1999, pp. 70 – 81, on-line la adresa:http:// ieeexplore. ieee.org/document/755469/ (accesată cel mai recent la 20 februarie 2019).

3. Garry Kasparov, Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins, Hodder & Stoughton, Londra, 2017.

4. The Good Knight, „Deep Blue vs Garry Kasparov Game 2 (1997 Match)", YouTube, urcat pe 18 octombrie 2012, accesibil on-line la adresa https://www.youtube.com/ watch? v = 3 BdlQ2 rOmok & t = 2290's (accesată cel mai recent la 20 februarie 2019).

5. Ibidem.

6. Steven Levy, „Big Blue's Hand of God", *Newsweek*, numărul din 18 mai 1997, on-line la adresa http://www.newsweek.com/big-blues-hand-god-173076 (accesată cel mai recent la 20 februarie 2019).

7. Garry Kasparov, op. cit., p. 187.

8. *Ibidem*, p. 191.

9. Definiţia din *Merriam-Webster*. The *Oxford English Dictionary's* subliniază bazele matematice ale algoritmilor: „un proces sau set de reguli care urmează să fie folosite în calcule matematice sau alte operaţii necesare în rezolvarea unei probleme oarecare, în principal, de către un computer".]

10. Există o multitudine de modalităţi diferite de grupare a algoritmilor şi nu mă îndoiesc că unii specialişti în programare se vor plânge că lista este (mult prea) simplificată. E adevărat că o listă exhaustivă ar fi inclus şi alte câteva categorii principale: algoritmi de mapare (*data mapping*, engl.), reducerea, regresia şi gruparea sau clasificarea datelor (iclustering, engl; *le partitionnement de données*, fr.), ca să nu enumăr decât câţiva. În cele din urmă, am ales acest tip anumit de categorii – apud

Nicholas Diakopoulos, *Algorithmic Accountability Reporting*: *On the Investigation of Black Boxes*, Tow Center for Digital Joumalism, Columbia University, New York, 2014, care acoperă foarte bine elementele de bază şi oferă o modalitate utilă de a clarifica şi de a condensa un domeniu de studiu vast şi complex.

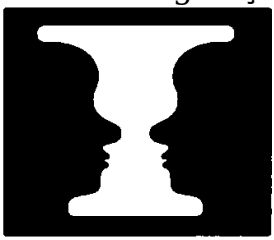
11. Kerbobotat, „Went to buy a baseball bat on Amazon, they have some interesting suggestions for accesories”, *Reddit*, postare din 28 septembrie 2013, on-line la adresa https://www.reddit.com/r/funny/comments/1nb 161/ wenttobuyabaseballbatonamazontheyhave/ (accesată cel mai recent la 20 februarie 2019).

12. Sarah Pérez, „Über debuts a «smarter» Über Pool în Manhattan”, *Techerunch*, 22 mai 2017, https://techerunch.com/2017/05/22/über-debuts-a-smarteruberpool-in-manhattan/.

13. Intenţionat am spus „teoretic”. Realitatea ar putea fi uşor diferită. Unii algoritmi au fost concepuţi, de-a lungul anilor, de către sute, chiar mii de programatori, fiecare adăugându-şi, treptat, propriii paşi în acest proces, însă, pe măsură ce liniile de cod se înmulţesc, complexitatea sistemului creşte, la rândul său, până când firele de execuţie (*threads*, engl.) ajung ca o porţie de spaghete încâlcite. În cele din urmă, algoritmul devine imposibil de urmărit şi mult prea complicat pentru ca vreo fiinţă umană să-l mai poată înţelege. În anul 2013, companiei Toyota i s-a cerut să plătească suma de trei milioane de dolari ca despăgubire în urma unui accident mortal în care a fost implicat unul dintre autovehiculele lor. Acesta accelerase incontrolabil, deşi conducătoarea auto apăsase pe frână, nicidecum pe acceleraţie în tot acel timp. Un specialist chemat să depună mărturie a declarat juriului că de vină era o instrucţiune accidentală ascunsă adânc în interiorul vastei încâlceli a software-ului. Vezi Phil Koopman, *A case study of Toyota unintended acceleration and software safety*, Pittsburgh: Carnegie Mellon University, prezentare creată la 18 septembrie 2014,

accesibilă Online la adresa https://users. ece. cmu.edu/-koopman/ pubs/koopmant 4 toyotauaslides.pdf (accesată cel mai recent la 20 februarie 2019).

14. Această iluzie optică (exemplul de aici este preluat de la adresa https://commons. wikimedia.org/wiki/File: Vaseof rubin.png, accesată cel mai recent la 20 februarie 2019) este cunoscută drept „Vaza lui Rubin", după Edgar Rubin, cel care a dezvoltat ideea. Este un exemplu de imagine *ambiguă* – situată chiar la limita dintre cele două profiluri închise la culoare care privesc unul spre celălalt şi imaginea unei vaze albe. Aşa cum apare desenată, este foarte uşor să treci de la una la alta din cele două în mintea ta şi nu ai nevoie decât de două linii din imagine pentru a bascula într-o direcţie şi cealaltă. Poate de conturul vag al ochilor în cazul feţelor sau de umbra care mărgineşte gâtul vazei.



O chestiune asemănătoare este şi exemplul de recunoaştere a imaginii câine/maşină. Echipa de cercetători a găsit o imagine care se situa chiar la limita dintre două clasificări diferite şi a folosit cea mai mică perturbare posibilă pentru a trece imaginea dintr-o categorie într-alta, din perspectiva calculatorului.

15. Jiawei Su, Danilo Vasconcellos Vargas şi Kouichi Sakurai, „One pixel attack for fooling deep neural networks", *arXiv. 1 719.08864 v4* [ce. LG], publicat la 22 februarie 2018, https://arxiv.org/pdf/1710.08864.pdf (accesat cel mai recent la 20 februarie 2019).

16. Chris Brooke, „«I was only following satnav orders» is no defence: driver who ended up teetering on cliff edge convicted of careless driving", *Daily Mail*, numărul din 16 septembrie 2009, on-line la adresa

http://www.dailymail.co.uk/news/article-1213891/Driver-ended-teetering-cliff-edge-guilty-blindly-following-sat-nav-directions. htmlixzz59 vihbQ2 n (accesată cel mai recent la 20 februarie 2019).

17. Ibidem.

18. Robert Epstein şi Ronald E. Robertson, „The search engine manipulation efect (SEME) and its possible impact on the outcomes of elections", *Proceedings of the National Academy of Sciences*, vol. 112, nr. 33/2015, pp. E4512 – 4521, on-line la adresa: http://www.pnas.org/content/112/33/ E4512 (accesată cel mai recent la 20 februarie 2019).

19. David Shultz, „Internet search engines may be influencing elections", *Science*, numărul din 7 august 2015, on-line la adresa: http://www.sciencemag.org/news/2015/08/internet-search-engines-may-be-influencing-elections (accesată cel mai recent la 21 februarie 2019).

20. Robert Epstein şi Ronald E. Robertson, *idem*.

21. Linda J. Cazul *KW versus Armstrong*, United States District Court for the District of Idaho, 2 mai 2012, proces a cărui hotărâre judecătorească poate fi citită on-line la adresa: https://scholar. google.co.uk/scholarcase? case = 1706216 8494596747089 & hl = en & assdt = 2006 (accesată cel mai recent la 21 februarie 2019) Burdick, „Accountability and automation bias", *International Journal of Human-Computer Studies*, vol. 52/2000, pp. 701 – 717, on-line la adresa: http://lskitka. people. uic.edu/IJHCS2.000.pdf (accesată cel mai recent la 21 februarie 2019).

11. Cazul *KW versus Armstrong* „United States District Court for the District of Idaho, 2 mai 2012, proces a cărui hotărâre judecătorească poate fi citită on-line la adresa: https://scholar. google.co.uk/scholarcase? case = 1706216849459 6747089 & hl = en & assdt = 2006 (accesată cel mai recent la 21 februarie 2019).

12. Jay Stanley, *Pitfalls of Artificial Intelligence*

Decision making Highlighted în Idaho ACLU Case, American Civil Liberties Union, 2 iunie 2017, on-line la adresa https://www.aclu.org/blog/privacy-technology/pitfalls-artificial-intelligence-decisionmaking-highlighted-idahoaclu-case (accesată cel mai recent la 21 februarie 2019).

13. „K.W. v. Armstrong", *Leagle*. *comy* publicat la 24 martie 2014, la adresa: https://www.leagle.com/decision/infdco20140326 c20 (accesată cel mai recent la 21 februarie 2019).

14. Ibidem.

15. ACLU Idaho staff, https://www.acluidaho.org/en/about/ staff (site accesat cel mai recent la 21 februarie 2019).

16. Jay Stanley, op. cit.

17. ACLU, *Ruling mandates important protections for* duc *process rights of Idahoans with developmental disabilities*, publicat la 30 martie 2016, on-line la adresa web: https://www.aclu.org/news/federal-court-rules-against-idaho-department-health-and-welfare-medicaid-class-action (accesată cel mai recent la 21 februarie 2019).

18. Jay Stanley, op. cit.

19. Ibidem.

20. *Ibid*.

21. *Ibid*.

22. *Ibid*.

23. Kristine Phillips, „The former Soviet officer who trusted his gut – and averted a global nuclear catastrophe", *Washington Post*, numărul din 18 septembrie 2017, on-line la adresa: https://www.Washingtonpost.com/news/retropolis/wp/2017/09/18/

the-former-soviet-officer-who-trus ted-his-gutand-averted-a-global-nuclear-catastrophe/? utm term =.6546 eOfobece (accesat cel mai recent la 21 februarie 2019).

35. Pavel Aksenov, „Stanislav Petrov: the man who may have saved the world", ştire şi articol BBC News, din

26 septembrie 2013, on-line la adresa http://www.bbe.co.uk/news/ world-europe-24280831 (link accesat cel mai recent la 21 februarie 2019).

36. Ibidem.

37. Stephen Flanagan, *Re*: *Accident at Smiler Rollercoaster*, *Alton Towers*, *2 June 2015*: *Expert's Report*, pregătit la cererea Direcţiei de Securitate şi Siguranţă, publicat online în octombrie 2015, la adresa http://www.chiark. greenend.org.uk/-ijackson/2016/Expert %20 witness%20 report%20 from%20 Steven%20 Flanagan.pdf (accesată cel mai recent la 21 februarie 2019).

38. Paul E. Meehl, *Clinical versus Statistical Prediction*: *A Teoretical Analysis and a Review of the Evidence*, University of Minnesota, Minneapolis, 1996, (după prima ediţie din 1954), on-line la adresa http://citeseerx. ist. psu.edu/viewdoc/download? doi = 10.1.1.693.6031 & rep = repl & type = pdf (accesată cel mai recent la 21 februarie 2019).

39. William M. Grove, David H. Zald, Boyd S. Lebow, Beth E. Snitz şi Chad Nelson, „Clinical versus mechanical prediction: a metaanalysis", *Psychological Assessment*, vol. 12, nr. 1/2000, p. 19.

40. Berkeley J. Dietvorst, Joseph P. Simmons şi Cade Massey.

„Algorithmic aversion: people erroneously avoid algorithms after seeing them err", *Journal of Experimental Psychology*, numărul din septembrie 2014, on-line la adresa: http://opim. wharton. upenn.edu/risk/library/

WPAF201410-Algorithm Aversion-Dietvorst-Simmons-Massey.pdf (accesată cel mai recent la 21 februarie 2019).

CAPITOLUL 2 – Datele personale informatizate

1. Nicholas Carlson, „Well, these new Zuckerberg IMS won't help Facebook's privacy problems", *Business Insider*, numărul din 13 mai 2010, care poate fi citit on-line http://www.businessinsider.com/well-these-newzuckerberg-ims-wont-help-facebooks-privacy-problems2010 – 5? IR = T

(link accesat cel mai recent la 6 martie 2019).

2. Clive Humby, Terry Hunt şi Tim Phillips, *Scoring Points*: *How Tesco Continues to Win Customer Loyalty*, Kogan Page, Londra, 2008.

3. *Ibidem*, ediţia Kindle, pp. 1313 – 1317.

4. See Eric Schmidt, „The creepy line", fragment de interviu urcat pe YouTube pe 11 februarie 2013, care poate fi vizionat on-line la adresa: https://www.youtube.com/watch?v=o-rvER6YTss (link accesat cel mai recent la 6 martie 2019).

5. Charles Duhigg, „How companies leant your secrets", *New York Times*, numărul din 16 februarie 2012, on-line la adresa: https://www.nytimes.com/2012/02/19/magazine/shoppinghabits.html (accesată cel mai recent la 6 martie 2019).

6. Ibidem.

7. Sarah Buhr, „Palantir has caised $880 million at a $20 billion valuation", *Techerunch*, numărul din 23 decembrie 2015.

8. Comisia Federală pentru Comerţ, *Data Brokers*: *A Call for Transparency and Accountability*, (Washington DC, mai 2014), on-line la adresa: https://www.ftc.gov/sistem/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may2014/140527 databrokerreport.pdf (accesată cel mai recent la 6 martie 2019).

9. Ibidem.

10. Wolfie Chrisd, *Corporate Surveillance în Everyday Life*, raport publicat de Institute for Criticai Digital Culture „Cracked Labs", Viena, iunie 2017, on-line la adresa http://crackediabs.org/en/corporate-surveillance (accesată cel mai recent la 6 martie 2019).

11. Heidi Waterhouse, „The death of data: retention, rot, and risk", conferinţă organizată de The Lead Developer, Austin, Texas, susţinută la 2 martie 2018, postată pe YouTube pe 21 martie 2018, on-line la adresa web: https://www.youtube.com/watch?v=mXvPCheo9 iU

(accesată cel mai recent la 6 martie 2019).

12. Amit Datta, Michael Carl Tschantz şi Anupam Datta, „Automated experiments on ad privacy settings”, *Proceedings on Privacy Enhancing Technologies*, nr. 1/2015, pp. 92 – 112.

13. Latanya Sweeney, „Discrimination in online ad delivery”, *in Queue*, vol. 11, nr. 3/2013, p. 10, on-line la adresa web: https://dl. acm.org/citation. din? id = 2460278 (accesată cel mai recent la 6 martie 2019).

14. Jon Brodkin, „Senate votes to let ISPS sell your Web browsing history to advertisers”, în *Ars Tehnica*, numărul din 23 martie 2017, on-line la adresa web: https://arstehnica.com/tech-policy/2017/03/senate-votes-to-let-isps-sell-your-web-browsing-history-to-advertisers/ (accesată cel mai recent la 6 martie 2019).

15. Svea Eckert şi Andreas Dewes, „Dark data”, DEFCON, Conferinţa 25 din 20 octombrie 2017, postată pe YouTube pe 20 octombrie 2017, on-line la adresa web: https://www.youtube.com/watch?v=invYGi7-Lxo (accesată cel mai recent la 7 martie 2019).

16. Cercetătorii s-au bazat în această parte a activităţii lor pe lucrarea lui Arvind Nārāyanan şi Vitaly Shmatikov, „Robust de-anonymization of large sparse datasets” prezentată în cadrul Simpozionului despre securitate şi confidenţialitate organizat de Institute of Electrical and Electronics Engineers (IEEE, Institutul Inginerilor Electrotehnişti şi Electronişti), în perioada 18 – 22 mai 2008.

17. Michal Kosiński, David Stillwell şi Thore Graepel. „Private traits and attributes are predictable from digital records of human behavior”, vol. 110, nr. 15/2013, pp. 5802 – 5805.

18. Ibidem.

19. Wu Youyou, Michal Kosiński şi David Stillwell, „Computer-based personality judgements are more accurate than those made by humans”, în *Proceedings of the National Academy of Sciences*, vol. 112, nr. 4/2015, pp.

1036 – 1040.

20. S.C. Matz, M. Kosiński, G. Nave şi DJ. Stillwell, „Psychological targeting as an efective approach to digital mass persuasion”, *Proceedings of the National Academy of Sciences*, vol. 114, nr. 48/2017, pp. 12714 – 12719, on-line la adresa https://www.pnas. Org/content/1 14/48/12714? ijk ey = 67 bf5 df8 e5307 e45 a26174 llb054 cbc508467668 & ke ytype2 = tfipsecsha (accesată cel mai recent pe 7 martie 2019).

21. Paul Lewis şi Paul Hilder, „Leaked: Cambridge Analytical blueprint for Trump victory”, publicat în *Guardian*, numărul din 23 martie 2018, on-line la adresa https://www.theguardian.com/uk-news/2018/mar/23/leaked -cambridgeanalyticas-blueprint-for-trump-victory (accesată cel mai recent pe 7 martie 2019).

22. „Cambridge Analytica planted fake news”, ştire BBC din 20 martie 2018, on-line la adresa: http://www.bbe.co.uk/ news/au/world-43472347/cambridge-analytica-plantedfake-news (accesată cel mai recent pe 7 martie 2019).

23. Adam D.I. Kramer, Jamie E. Guillory şi Jeffrey T. Hancock, „Experimental evidence of massive-scale emoţional contagion through social networks”, *Proceedings of the National Academy of Sciences*, vol. III, nr. 24/2014, pp. 8788 – 8790, on-line la adresa web: https://www.pnas.org/ content/111/24/8788 (accesată cel mai recent pe 7 martie 2019).

24. Jamie Bartlett, „Big data is watching you – and it wants your vote”, în *Spectator*, numărul din 24 martie 2018, on-line la adresa web: https://www.spectator.co.uk/2018/03/big-data-is-watching-you-and-it-wants-your-vote/ (accesată cel mai recent pe 7 martie 2019).

25. Li Xiaoxiao, „Ani Financial Subsidiary Starts Offering Individual Credit Scores”, în *Caixin*, articol postat pe

2 martie 2015, on-line la adresa:

https://www.caixinglobal.com/2015 – 03 – 02/101012655.html (accesată cel mai recent pe 7 martie 2019).

26. Rick Falkvinge, „În China, your credit score is now affected by your political opinions – and your friends political opinions", în *Privacy News Online*, postat pe 3 octombrie 2015, on-line la adresa web: https://www.privateinternetaccess.com/blog/2015/10/inchina-your-credit-score-isnow-affected-by-your-political-opinions-and-your-friends-political-opinions/ (accesată cel mai recent pe 7 martie 2019).

27. Rogier Criemers, *State Council Guiding Opinions Concerning Establishing and Perfecting Incentives for Promisekeeping and Joint Punishment Sistems for Trust-breaking, and Accelerating the Construction of Social Sincerity*, articol publicat pe blogul China Copyright and Media la 30 mai 2016, actualizat la 18 octombrie 2016, on-line la adresa web: https://chinacopyrightandmedia.wordpress.com/2016/05/30/ state-council-guiding-opinions-concerning-estabilishing-and-perfecting-incentives-for-promisekeeping-and-joint-punishment-systems-for-trust-breaking-and-accelerating-the-construcţion-of-social-sincer/ (accesată cel mai recent la 7 martie 2019).

28. Rachel Botsman, Who Can You Trust? How Technology Brought Us Together – and Why It Could Drive Us Apart, Penguin, Londra, 2017, apud ediţia Kindle, p. 151.

CAPITOLUL 3 – *Justiţia*

1. John-Paul Ford Rojas, „London riots: Lidl water thief jailed for six months", articol în *Telegraph*, numărul din 7 ianuarie 2018, on-line la adresa http://www.telegraph.co.uk/news/uknews/crime/8695988/London-riots-Lidl-waterthief-jailed-for-six-months.html (accesară cel mai recent la 22 martie 2019).

2. Matthew Taylor, „London riots: how a peaceful

festival în Brixton turned into a looting free-for-all", în *Guardian*, numărul din 8 august 2011, on-line la adresa https://www.theguardian.com/uk/2011/aug/08/london-riots-festivalbrixton-looting (accesată cel mai recent la 22 martie 2019).

1. John-Paul Ford Rojas, op. cit.

2. Josh Halliday, „London riots: how Blackberry Messenger played a key role", în *Guardian*, numărul din 8 august 2011, on-line la adresa https://www.theguardian.com/media/ 201 l/aug/08/london-riots-facebook-twitter-blackberry (accesată cel mai recent la 22 martie 2019).

3. David Mills, „Paul and Richard Johnson avoid prison over riots", în *News Shopper*, numărul din 13 ianuarie 2012, on-line la adresa http://www.newsshopper.co.uk/londormots/9471288. Fatherandsonavoidprisonoverriots/ (accesată cel mai recent la 22 martie 2019).

4. Ibidem.

5. John-Paul Ford Rojas, op. cit. „În mod normal, poliţia nu te-ar aresta pentru acest gen de infracţiune. Nu te-ar ţine în arest. Nu te-ar trimite în judecată" a scris, în 2015, Hannah Quirk, conferenţiar universitar la catedra de justiţie şi drept penal a Universităţii din Manchester despre cazul lui Nicholas, apud. Carly Lightowlers şi Hannah Quirk, „The 2011 English „riots": prosecutorial zeal and judiciar abandon", în *British Journal of Criminology*, vol. 55, nr. 1/2015, pp. 65 – 85.

6. David Mills, op. cit.

7. William Austin şi Thomas A. Williams III, „A survey of judges responses to simulated legal cases: research note on sentencing disparity", în *Journal of Criminal Law and Criminology*, vol. 68, nr. 2/1977, pp. 306 – 310.

8. Mandeep K. Dhami şi Peter Ayton, „Bailing and jailing the fast and frugal way", *Journal of Behavioral Decision-making*, în vol. 14, nr. 2/2001, pp. 141 – 68, on-line la adresa: http://onlinelibrary.

wiley.com/doi/10.1002/bdm.371 /abstract (accesată cel mai recent la 22 martie).

9. Aproape jumătate dintre judecători au avut opinii diferite despre cel mai bun curs al acţiunii în fiecare dintre cazuri.

12. Statisticienii au o metodă de măsurare a acestui gen de coerenţă a hotărârilor, numită „coeficientul lui Cohen" (kappa). Acesta ia în considerare faptul că (până şi în cazul în care ai face o presupunere absolut la întâmplare) ai putea avea norocul să nimereşti exact. Valoarea unu înseamnă o coerenţă perfectă. Valoarea zero e ca şi cum ai fi ghicit la întâmplare. Valorile obţinute de judecători au fost cuprinse între zero şi unu, cu o medie de 0,69.

13. Diane Machin, „Sentencing guidelines around the world", o lucrare pregătită pentru Scottish Sentencing Council, în mai 2005, on-line la adresa: https://www.scottishsentencingcouncil.org.uk/media/1109/ paper-31 a-sentencingguidelines-around-the-world.pdf (accesată cel mai recent la 22 martie 2019).

14. Ibidem.

15. Ibid.

16. Ernest W. Burgess, „Factors determining succes or failure on parole", în *The Workings of the Intermediate-sentence Law and Parole Sistem în Illinois* (State Board of Parole, Springfield, Illinois, 1928). Este o lucrare greu de găsit, aşa că vă sugerez o lectură alternativă care îi aparţine colegului lui Burgess, Tibbitts, despre un studiu ulterior celui original, „Succes or failure on parole can be predicted: a study of the records of 3.000 youths paroled from the Illinois State Reformatory", în *Journal of Criminal Law and Criminology*, vol. 22, nr. 1/primăvara 1931, on-line la adresa https://scholarlycommons. law. northwestem.edu/cgi/ viewcontent. cgi? article = 2211 & context = jeic (accesată cel mai recent la 22 martie 2019). Celelalte categorii utilizate de Burgess erau „oaie neagră", „infractor din întâmplare", „greu de cap" şi „bandit". El a

descoperit că aceia din categoria „băiat de la fermă” erau cel mai puţin predispuşi să recidiveze.

17. Karl F. Schuessler, „Parole prediction: its history and status”, în *Journal of Criminal Law and Criminology*, vol. 45, nr. 4/1955, pp. 425 – 31, on-line la adresa: https://

pdfs. semanticscholar.org/4 cd2/31 dd25321 aOcl 4 a93 58 a93 ebecbbfl5 d3169.pdf (accesată cel mai recent la 22 martie 2019).

18. Ibidem.

19. Bernard E. Harcourt, *Against Prediction*: *Profiling*, *Policing*, *and Punishing în an Actuarial Age*, University of Chicago Press, Chicago şi Londra, 2007, p. 1.

20. Philip Howard, Brian Francis, Keith Soothill şi Les Humphreys, *OGRS 3*: *The Revised Offender Group Reconviction Scale*, Sinteza cercetării 7/09, Ministry of Justice (Ministerul de justiţie britanic), Londra, 2009, on-line la adresa https://core.ac.uk/download/pdf/1556521.pdf (accesată cel mai recent la 22 martie 2019).

21. Se cuvine să fac un mic avertisment aici: în această statistică este posibil să existe o eroare de selecţie. „Întreabă publicul” se folosea în mod obişnuit în primele etape ale jocului, când întrebările erau mult mai uşoare. Însă ideea că opiniile colective ale unui grup sunt mai corecte decât cele ale unei singure persoane, reprezintă un fenomen bine documentat. Pentru mai multe informaţii despre acest subiect, vezi James Surowiecki, *The Wisdom of Crowds*: *Why the Many Are Smarter than the Few*, Doubleday, New York, 2004, p. 4.

22. Netflix Technology Blog, on-line la adresa https://medium.com/netflix-techblog/netflix-recommendations-beyondthe-5-stars-part-2-d9 b96 aa399 f5 (accesată cel mai recent la 22 martie 2019).

23. Shih-ho Cheng, „Unboxing the random forest classifier: the threshold distributions”, Airbnb Engineering and Data Science, on-line la adresa: https://medium.com/airbnbengineering/unboxing-the-random-forest-classifier-thethreshold-distributions-22

ezb58 ea6 (accesată cel mai recent la 22 martie 2019).

24. Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig şi Sendhil Mullainathan, *Human Decisions and Machine Predictions*, Documentul de lucru nr. 23180 al

National Bureau of Economic Research (NBER, Biroul Naţional de Cercetări Economice, ONG american), National Bureau of Economic Research, Cambridge, Massachussetta, februarie 2017), on-line la adresa: http://www.nber.org/ papers/w23180 (accesată cel mai recent la 22 martie 2019). Studiul acesta a folosit, de fapt „arbori de decizie de tip *gradient boosted*", un algoritm asemănător cu *random forests*. Ambii combină predicţiile unei multitudini de arbori de decizie pentru pentru a ajunge la un rezultat, însă arborii din metoda *gradient-boosted* cresc secvenţial, pe când în *random forests* cresc în paralel. Pentru realizarea acestui studiu, setul de date a fost, mai întâi, înjumătăţit. O jumătate a fost folosită pentru antrenarea algoritmului, cealaltă jumătate a fost păstrată deoparte. În momentul în care algoritmul a fost pregătit, a luat cazuri din jumătatea pe care nu o văzuse niciodată şi a încercat să prevadă ce avea să se întâmple. (Dacă nu s-ar împărţi, mai întâi, informaţiile, algoritmul nu ar fi decât un frumos tabel de căutare.)

25. Oamenii de ştiinţă au petrecut foarte mult timp dezvoltând tehnici statistice care să se ocupe exact de această problemă, astfel încât să se poată face, în continuare, o comparaţie semnificativă între predicţiile făcute de judecători şi, respectiv, de algoritmi. Pentru mai multe informaţii legate de această temă, vezi Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig şi Sendhil Mullainathan, op. cit.

26. „Costs per place and costs per prisoner by individual prison", *National Offender Management Service Annual Report and Accounts 2015 – 16*, Act adiţional de gestionare a informaţiilor din Comunicatul Ministerului de Justiţie, 27 octombrie 2016, on-line la adresa:

https://www.gov.uk/government/uploads/sistem/uploads/attachmentdata/file/563326/costs-per-place-cost-per-prisoner-2015 – 16.pdf (accesată cel mai recent la 22 martie 2019).

27. Marc Santora, „City's annual cost per înmate is $168.000, study finds", în *New York Times*, 23 august 2013, on-line la adresa: http://www.nytimes.com/2013/08/24/nyregion/ citys-annual-cost-per-inmate-is-nearly-168000-studysays.html (accesată cel mai recent la 22 martie 2019); Universitatea Harvard, „Harvard at a glance", on-line la adresa https://www.harvard.edu/about-harvard/harvard-glance (accesată cel mai recent la 22 martie 2019).

28. Luke Dormehl, The Formula: HowAlgorithms Solve AU Our Problems... and Create More, W.H. Allen, Londra, 2014, p. 123.

29. Julia Angwin, Jeff Larson, Surya Mattu şi Lauren Kirchner, „Machine bias", în ProPublica, 23 mai 2016, on-line la adresa https://www.propublica.org/article/machinebias-risk-assessments-in-criminal-sentencing (accesată cel mai recent la 22 martie 2019).

30. Chestionar de „Evaluare a riscurilor", on-line la adresa https://www.documenteloud.org/documents/2702103 – Sample-Risk-Assessment – COMPAS-CORE.html (accesată cel mai recent la 22 martie 2019).

31. Tim Brennan, William Dieterich şi Beate Ehret (Institutul Northpointe), „Evaluating the predictive validity of the COMPAS risk and needs assessment sistem", în *Criminal Justice and Behavior*, vol. 36, nr. 1/2009, pp. 21 – 40, on-line la adresa http://www.northpointeinc.com/files/ publications/Criminal-Justice-Behavior – COMPAS.pdf (accesată cel mai recent la 22 martie 2019). Conform unui studiu din 2018, algoritmul COMPAS are o precizie asemănătoare cu un „ansamblu" de persoane. Cercetătorii au demonstrat că, dacă cereau unui grup de 20 de

persoane fără experienţă să prezică recidiva, obţineau un punctaj echivalent cu acela al sistemului COMPAS. Este o comparaţie interesantă, însă merită să ne amintim că în sălile de judecată adevărate, judecătorii nu dispun de un grup de persoane necunoscute care să voteze în camera din spate. Sunt pe cont propriu. Iar aceasta este singura comparaţie care contează cu adevărat. Vezi Julia Dressel şi Hany Farid, „The accuracy, fairness, and limits of predicting recidivism", în *Science Advances*, vol. 4, nr. 1/2018.

32. *Christopher Drew Brooks* v. *Commonwealth*, Curtea de Apel a statului Virginia, Memorandum de acord al judecătorului Rudolph Bumgardner III, 28 ianuarie 2004, on-line la adresa https://law.justia.com/cases/Virginia/court-ofappeals-unpublished/2004/2540023.html (accesată cel mai recent la 22 martie 2019).

33. „ACLU brief challenges constitutionality of Virginia's sex offender risk assessment guidelines", American Civil Liberties Union Virginia, 28 octombrie 2003, on-line la adresa https://acluva.org/en/press-releases/adu-briefehallenges-constitutionality-virginias-sex-offender-riskassessment (accesată cel mai recent la 22 martie 2019).

34. *State* v. *Loomis*, Curtea Supremă a statului Wisconsin, 13 iulie 2016, on-line la adresa http://caselaw.findiaw.com/ wi-supreme-court/1742124.html (accesată cel mai recent la 22 martie 2019).

35. Citatele provin dintr-o discuţie particulară cu Richard Berk.

36. Julia Angwin, Jeff Larson, Surya Mattu şi Lauren Kirchner, op. cit.

37. *Global Study on Homicide 2013*, Biroul ONU pentru Droguri şi Criminalitate, Viena, 2014, on-line la adresa http://www.unode.org/documents/gsh/pdfs/2014GLOBAL HOMICIDEBOOK – web.pdf (accesată cel mai recent la 22

martie 2019).

38. ACLU, „The war on marijuana în black and white”, iunie 2013, on-line la adresa www.aclu.org/files/assets/acluthe waronmarijuana-vel2.pdf (accesată cel mai recent la 22 martie 2019).

39. În mod surprinzător, poate, poziţia companiei Equivant asupra acestui aspect este susţinută de Curtea Supremă a Statului Wisconsin. După ce Eric Loomis a fost condamnat la şase ani de închisoare de către un judecător care a utilizat instrumentul COMPAS de evaluare a riscului, el a făcut apel împotriva acestei sentinţe. În procesul *Loomis* vs. *Wisconsin*, s-a susţinut că utilizarea unui software brevetat, cu circuit închis, de evaluare a riscului pentru stabilirea sentinţei sale îi încălca dreptul la un proces echitabil, întrucât apărarea nu putea să pună sub semnul întrebării validitatea ştiinţifică a punctajului obţinut. Însă Curtea Supremă a Statului Wisconsin a stabilit că utilizarea în instanţă, cu titlu de încercare, a unei evaluări algoritmice a riscului în luarea deciziei nu încălca dreptul inculpatului la un proces echitabil: Curtea Supremă a Statului Wisconsin, cazul nr. 2015AP157-CR, verdict înregistrat la 13 iulie 2016, on-line la adresa https://www.wicourts.gov/se/opinion/ Display Document.pdf? content = pdf & seq No = 171690 (accesată cel mai recent la 22 martie 2019).

40. Lucy Ward, „Why are there so few female maths profesors în universities?”, *Guardian*, 11 martie 2013, on-line la adresa https://www.theguardian.com/lifeandstyle/thewomens-blog-with-jane-martinson/2013/mar/11 /women-maths-professors-uk-universities (accesată cel mai recent la 22 martie 2019).

41. Sonja B. Starr şi M. Mărit Rehavi, *Racial Disparity în Federal Criminal Charging and Its Sentencing Conséquences*, Program de Drept şi Ştiinţe Economice, Document de lucru nr. 12 – 002, Universitatea de Drept din Statul Michigan, Ann Arbor, 7 mai 2012, on-line la adresa

http://economics. ube.ca/files/2013/05/pdfpapermarit-rehavi-racialdisparity-federal-criminalpdf (accesată cel mai recent la 22 martie 2019).

42. David Arnold, Will Dobbie şi Crystal S. Yang, *Racial Bias* în *Bail Decisions*, NBER Document de lucru nr. 23421, Biroul Naţional de Cercetări Economice – NBER, Cambridge, Massachusetts, 2017, on-line la adresa https://www.princeton.edu/-wdobbie/files/racialbias.pdf (accesată cel mai recent la 22 martie 2019).

43. John J. Donohue III, Capital Punishment în Connecticut, 1973 – 2007: A Comprehensive Evaluation from 4686

Murders to One Execution, Şcoala de Drept Stanford şi Biroul Naţional de Cercetări Economice, Stanford, California şi Cambridge, Massachusetts, octombrie 2011, on-line la adresa https://law. stanford.edu/wp-content/ uploads/sites/default/files/publication/259986/doc/ şi spublic/fulltext.pdf (accesată cel mai recent la 22 martie 2019).

44. Adam Benforado, *Unfair*: *The New Science of Criminal Injustice*, Crown, New York, 2015, p. 197.

45. Sonja B. Starr, *Estimating Gender Disparities în Federal Criminal Cases*, Universitatea Michigan, Facultatea de Drept şi Articol ştiinţific pe teme de economie nr. 12 – 018, Universitatea de Drept din Statul Michigan, Ann Arbor, 29 august 2012, on-line la adresa https://ssm.com/ abstract = 2144002 ori la http://dx. doi.org/10.2139/ ssm.2144002 (accesată cel mai recent la 22 martie 2019).

46. David B. Mustard, „Racial, ethnic, and gender disparities în sentencing: evidence from the US federal courts", *Journal of Law and Economics*, vol. 44, nr. 2/aprilie 2001, pp. 285 – 314, on-line la adresa http://people. terry. uga.edu/mustard/sentencing.pdf (accesată cel mai recent la 22 martie 2019).

47. Daniel Kahneman, *Thinking, Fast and Slow*, Farrar, Straus & Giroux, New York, 2011, p. 44 [*Gândire*

rapidă, *gândire lentă*, traducere în limba română de Dan Crăciun, Editura Publica, Bucureşti, 2015].

48. Chris Guthrie, Jeffrey J. Rachliński şi Andrew J. Wistrich, *Blinking on the Bench*: *How Judges Decide Cases*, lucrarea nr. 917, Facultatea de Drept a Universităţii Cornell, New York, 2007, on-line la adresa http://scholarship. law. cornell.edu/fac pub/917 (accesată cel mai recent la 22 martie 2019).

49. Daniel Kahneman, op. cit., p. 13.

50. *Ibidem.*, p. 415.

51. Mandeep K. Dhami şi Peter Ayton, op. cit.

44. Brian Wansink, Robert J. Kent şi Stephen J. Hoch, „An anchoring and adjustment model of purchase quantity decisions", *Journal of Marketing Research*, vol. 35, 1998, pp. 71 – 81,http://foodpsychology. cornell.edu/sites/default/files/unmanagedfiles/AnchoringMR-1998.pdf (accesată cel mai recent la 22 martie 2019).

45. Mollie Marţi şi Roselle Wissler, „Be careful what you ask for: the efect of anchors on personal înjury damages awards", *Journal of Experimental Psychology*: *Applied*, vol. 6, nr. 2/2000, pp. 91 – 103.

46. Bine Englich şi Thomas Mussweiler, „Sentencing under uncertainty: anchoring effects in the courtroom", *Journal of Applied Social Psychology*, vol. 31, nr. 7/2001, pp. 1535 – 1551, on-line la adresa http://onlinelibrary. wiley.com/doi/10. llll/j.1559 – 1816.2001. th02687. x (accesată cel mai recent la 22 martie 2019).

47. Bine Englich, Thomas Mussweiler şi Fritz Strack, „Playing dice with criminal sentences: the influence of irrelevant anchors on experts judiciar decision making", *Personality and Social Psychology Bulletinf* vol. 32/2006, pp. 188 – 200, on-line la adresa https://www.researchgate.net/publication/7389517 Playing Dice Withiriminal Sentences ThelnfluenceoflrrelevantănchorsonexpertS%27 Judicial Decision Making? enrichld = rgreqzedfeb71 aa83 f8 fad80 cc24 df3254 d-XXX & enrichăourc e =

Y292ZXJQYWdlOzezODklMTc7QVM6MTAzODIzNjlw MTgyMDIyQDEOMDE3NjQ4ODgzMTA%3D & el = lx3 & esc = publicaţionlover Pdf (accesată cel mai recent la 22 martie 2019).

48. Ibidem.

49. Ibid.

50. Mandeep K. Dhami, Ian K. Belton, Elizabeth Merrall, Andrew McGrath şi Sheila Bird, „Sentencing în doses: is individualizând justice a myth?", în curs de revizuire. Comunicată cu amabilitate într-o discuţie particulară cu Mandeep Dhami.

51. Ibidem.

52. Adam N. Glynn şi Maya Sen, „Identifying judiciar empathy: does having daughters cause judges to rule for women's issues?", *American Journal of Political Science*, vol. 59, nr. 1/2015, pp. 37 – 54, on-line la adresa https://scholar. harvard.edu/files/msen/files/daughters.pdf (accesată cel mai recent la 22 martie 2019).

53. Shai Danziger, Jonathan Levav şi Liora Avnaim-Pesso, „Extraneous factors în judiciar decisions", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, nr. 17/2011, pp. 6889 – 6892, on-line la adresa http://www.pnas.org/content/108/17/6889 (accesată cel mai recent la 22 martie 2019).

54. Keren Weinshall-Margel şi John Shapard, „Overlooked factors in the analysis of parole decisions", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, nr. 42, /2011, E833, on-line la adresa http://www.pnas.org/content/108/42/E833. long (accesată cel mai recent la 22 martie 2019).

55. Un Simonsohn şi Francesca Gino, „Daily horizons: evidence of narrow bracketing în judgment from 9.000 MBA admission interviews", *Psychological Science*, vol. 24, nr. 2/2013, pp. 219 – 224, on-line la adresa https://ssm.com/ abstract = 2070623 (accesată cel mai recent la 22 martie 2019).

56. Lawrence E. Williams şi John A. Bargh, „Experiencing physical warmth promotes interpersonal warmth”, *Science*, vol. 322, nr. 5901, pp. 606 – 607, on-line la adresa https://www.nebi. nim. nih.gov/pâne/articles/PMC2737341/ (accesată cel mai recent la 22 martie 2019).

CAPITOLUL 4 – *Medicina*

1. Richard M. Levenson, Elizabeth A. Kmpiński, Victor M. Navarro şi Edward A. Wasserman. „Pigeons (*Columba livra*) as trainable observers of pathology and radiology breast cancer images”, *PLOSOne*, 18 noiembrie 2015, on-line la adresa http://joumals. plos.org/plosone/ article? id = 10.1371/joumal. pone. Ol 41357 (accesată cel mai recent la 22 martie 2019).

„Hippocrates daughter as a dragon kills a knight, în „The Travels of Sir John Mandeville”, *British Library Online Galleryy* 26 martie 2009, on-line la adresa http://www.al.uk/onlinegallery/onlineex/illmanus/harimanu coll/ h/01 Ihrl000003954 u00008 v00.html (accesată cel mai recent la 22 martie 2019).

Eleni Tsiompanou, „Hippocrates: tuneless still”, JLL Bulletin: Comentaries on the History of Treatment Evaluation, James Lind Library, Oxford şi Edinburgh, 2012, online la adresa http://www.jameslindlibrary.org/articles/ hippocrates-timeless-still/ (accesată cel mai recent la 22 martie 2019).

David K. Osborne, „Hippocrates: father of medicine”, *Greek Medicine*. *nety* 2015, on-line la adresa http://www.greekmedicine.net/whoswho/Hippocrates.html (accesată cel mai recent la 22 martie 2019).

Richard Coloan, „Is there room for art în evidence-based medicine?”, *AMA Journal of Ethics*, Virtual Mentor 13: 1, ianuarie 2011, pp. 52 – 54, on-line la adresa http://joumalofethics. ama-assn.org/2011/01/msocl-l 101.html. (accesată cel mai recent la 22 martie 2019)

Joseph Needham, *Science and Civilization in China*, vol. 6, *Biology and Biological Technology*, partea a VI-a,

Mediciney Nathan Sivin (editor), Cambridge University Press, Cambridge, 2004, p. 143, on-line la adresa https://monoskop. Org/images/l/16/Needhamlosephăcience andlivilisationinihinavol6 – 6 Biologyandbiological Technology Medicine.pdf (accesată cel mai recent la 22 martie 2019).

„Ignaz Semmelweis", *Brought to Life*: *Exploring the History of Medicine*, Muzeul Ştiinţei, Londra, f.a., on-line la adresa http://broughttolife. sciencemuseum.org.uk/

broughttolife/peoplezignazsemmelweis (accesată cel mai recent la 22 martie 2019).

8. Citatele provin dintr-o discuţie particulară purtată cu Andy Beck.

9. Joann G. Elmore, Gary M. Longton, Patricia A. Carney, Berta M. Geller, Tracy Onega, Anna N.A. Tosteson, Heidi D. Nelson, Margaret S. Pepe, Kimberly H. Allison, Stuart J. Schnitt, Frances P. O'Malley şi Donald L. Weaver, „Diagnostic concordance among pathologists interpreting breast biopsy specimens", *Journal of the* American *Medical Association*, vol. 313, nr. 11/17 martie 2015, 1122 – 1132, on-line la adresa https://jamanetwork.com/joumals/ jama/fullarticle/2203798 (accesată cel mai recent la 22 martie 2019).

10. Ibidem.

11. Numele de „reţea neurală" a apărut ca o analogie la ce se întâmplă în creier. Acolo, miliarde de neuroni sunt conectaţi între ei, într-o reţea gigantică. Fiecare neuron reacţionează la propriile conexiuni şi emite un semnal atunci când detectează starea de excitare a altui neuron. Semnalul duce apoi la excitarea altor neuroni care îl recepţionează. O reţea neurală este o versiune mult simplificată şi mult mai ordonată a creierului. Neuronii ei (artificiali) sunt structuraţi pe niveluri şi toţi neuronii de la un anumit nivel ascultă de toţi neuronii de la nivelul anterior. În exemplul cu câinele, primul şi cel dintâi nivel este constituit din pixelii individuali din imagine. Mai sunt apoi câteva niveluri având mii de neuroni în ele şi un nivel

final, care nu are decât un singur neuron de ieşire capabil să decidă dacă imaginea introdusă în reţea reprezintă un câine. Procedura de updatare a neuronilor este cunoscută drept „algoritm *backpropagation*”. Începem cu neuronul final, de ieşire, care decide probabilitatea ca imaginea introdusă în reţea să reprezinte un câine. Să spunem că am introdus imaginea unui câine, iar el a prezis că acea imagine avea şanse de 70% să reprezinte un câine. Neuronul de ieşire ia în considerare semnalele primite de la nivelul anterior şi spune: „Data viitoare când mai primesc o informaţie ca aceasta, voi majora probabilitatea ca imaginea să reprezinte un câine”. Apoi îi comunică fiecărui neuron de pe nivelul anterior: „Hei, dacă mi-aţi fi dat semnalul *acesta*, aş fi făcut o predicţie mai bună”. Fiecare dintre acei neuroni îşi analizează semnalele de intrare şi îşi modifică semnalul de ieşire pentru data viitoare. Apoi comunică nivelului anterior ce semnale trebuia să fi trimis şi aşa mai departe prin toate nivelurile, până când se ajunge la cel dintâi. Acest proces de propagare înapoi a erorilor prin reţeaua neurală i-a dat numele de „algoritm bacfepropagation” (de „propagare înapoi”). Pentru o perspectivă mai detaliată despre reţelele neurale, despre cum sunt construite şi antrenate acestea, vezi Pedro Domingos, *The Master Algorithm*: *How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, New York, 2015.

12. Alex Krizhevsky, Ilya Sutskever şi Geoffrey E. Hinton, „Image Net classification with deep convoluţional neural networks”, în F. Pereira, C.J.C. Burges, L. Bottou şi K.Q. Weinberger (editori), *Advances în Neural Information Processing Sistems 25*, Neural Information Processing Sistems Foundation, La Jolla, California, 2012, pp. 1097 – 1105, on-line la adresa http://papers. nips.cc/paper/4824-imagenet-classification-with-deep-convoluţional-neural-networks.pdf (accesată cel mai recent la 22 martie 2019). Acest tip de algoritmi este cunoscut sub numele de „reţea neurală de convoluţie”. În loc de introducerea unei imagini

întregi, algoritmul aplică, mai întâi, o serie de filtre şi caută tipare locale din modul în care este distorsionată imaginea.

13. Marco Tulio Ribeiro, Sameer Singh şi Carlos Guestrin, „«Why should I trust you?» Explaining the predictions of any classifier”, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 2016, pp. 1135 – 1144, on-line la adresa http://www.kdd.org/kdd2016/papers/ files/rfp0573-ribeiroA.pdf (accesată cel mai recent la 22 martie 2019).

14. Acest rezultat a fost comparat cu evaluarea făcută de mai mulţi specialişti, a căror analiză colectivă a fost considerată ca reprezentând un „adevăr bazat pe observaţie, indubitabil empiric” în privinţa conţinutului lamelelor.

15. Trafton Drew, Melissa L.H. Vo şi Jeremy M. Wolfe, „The invisible gorilla strikes again: sustained inattenţional blindness în expert observers”, *Psychological Science*, vol. 24, nr. 9/septembrie 2013, pp. 1848 – 1853, on-line la adresa https://www.nebi. nim. nih.gov/pâne/articles/ PMC3964612/ (accesată cel mai recent la 22 martie 2019).

16. Gorila este poziţionată în partea dreaptă sus a imaginii.

17. Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E. Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q. Nelson, Greg S. Corrado, Jason D. Hipp, Lily Peng şi Martin C. Stumpe, „Detecting cancer metastases on gigapixel pathology images”, Biblioteca Universităţii Cornell, 8 martie 2017, on-line la adresa https://arxiv.org/abs/1703.02442 (accesată cel mai recent la 22 martie 2019).

18. Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad şi Andrew H. Beck, „Deep learning for identifying metastatic breast cancer”, Biblioteca Universităţii Cornell, 18 iunie 2016, on-line http’s:

//arxiv.org/abs/1606.05718 (accesată cel mai recent la 22 martie 2019).

19. David A. Snowdon, „The Nun Study”, *Boletín de LAZOS de la Asociación Alzheimer de Monterrey*, vol. 4, nr. 22/2000; D.A. Snowdon, „Healthy aging and dementia: findings from the Nun Study”, *Annals of Internal Medicine*, vol. 139, nr. 5/septembrie 2003, pp. 450 – 454.

20. Densitatea ideilor – un substitut pentru complexitatea lingvistică – a fost calculată prin numărarea ideilor unice folosite de fiecare călugăriţă într-un şir de zece cuvinte. Aveţi o frumoasă prezentare aici: Associated

Press, „Study of nuns links early verbal skills to Alzheimer’s”, *Los Angeles Times*, 21 februarie 1996, on-line la adresa http://articles. latimes.com/1996 – 02 – 21/ news/mn-38356 lalzheimer-nuns-studied (accesată cel mai recent la 22 martie 2019).

21. Maya Nielsen, Jom Jensen şi Johan Andersen, „Precancerous and cancerous breast lesions during lifetime and at autopsy: a study of 83 women”, *Cancer*, vol. 54, nr. 4/1984, pp. 612 – 615, on-line la adresa http://onlinelibrary. wiley.com/ woll/doi/10.1002/1097 – 0142 (1984) 54: 4%3C612: AID-CNCR2820540403%3E3.O. CO; 2-B/abstract (accesată cel mai recent la 22 martie 2019).

22. H. Gilbert Welch şi William C. Black, „Using autopsy series to estimate the disease „rezervoir” for ductal carcinoma in situ of the breast: how much more breast cancer can we fiind?”, *Annals of Internal Medicine*, vol. 127, nr. 11/decembrie 1997, pp. 1023 – 1028, on-line la adresa www.vaouteomes.org/papers/Autopsy Series.pdf (accesată cel mai recent la 22 martie 2019).

23. Obţinerea unei statistici exacte este dificilă, pentru că ea depinde de ţară şi de demografie (precum şi de intensitatea cu care o anumită ţară face examinări pentru detectarea cancerului mamar). Pentru un rezumat de calitate, vezi: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/

breast-cancer (accesată cel mai recent la 22 martie 2019).

24. Citatele provin dintr-o discuţie particulară purtată cu Jonathan Kanevsky.

25. „Breakthrough method predicts risk of DCIS becoming invazive breast cancer”, *Artemis*, mai 2010, on-line la adresa http://www.hopkinsbreasteenter. Org/artemis/201005/3.html (accesată cel mai recent la 22 martie 2019).

26. H. Gilbert Welch, Philip C. Prorok, A. James O’Malley şi Barnett S. Kramer, „Breast-cancer tumor size, overdiagnosis, and mammography screening efectiveness”, *New England Journal of Medicine*, vol. 375/2016, pp. 1438 – 1447, on-line la adresa http://www.nejm.org/doi/full/10.1056/ NEJM – oal600249 (accesată cel mai recent la 22 martie 2019).

27. Un grup independent din Marea Britanie despre depistarea cancerului la sân, „The benefits and harms of breast cancer screening: an independent review”, *Lancet*, vol. 380, nr. 9855/30 octombrie 2012, pp. 1778 – 1786, on-line la adresa http://www.thelancet.com/joumals/lancet/article/ PIIS0140 – 6736 (12) 61611-O/abstract (accesată cel mai recent la 22 martie 2019).

28. Informaţii obţinute în timpul unei discuţii cu Jonathan Kanevsky.

29. Andrew H. Beck, Ankur R. Sangoi, Samuel Leung, Robert J. Marinelli, Torsten O. Nielsen, Marc J. van de Vijver, Robert B. West, Matt van de Rijn şi Daphne Roller, „Sistematic analysis of breast cancer morphology un covers stromal features associated with survival”, *Science Transitional Medicine*, 19 decembrie 2014, on-line la adresa https://becklab. hms. harvard.edu/files/becklab/files/scitransimed-2011 – beck108 ral 13.pdf (accesată cel mai recent la 22 martie 2019).

30. Phi Vu Tran, „A fully convoluţional neural network for cardiac segmentation in short-axis MRI”, 27

aprilie 2017, online la adresa https://arxiv.org/pdf/1604.00494.pdf (accesată cel mai recent la 22 martie 2019).

31. „Emphysema”, *Imaging Analytics*, Zebra Medical, on-line la adresa https://www.zebra-med.com/algorithms/lungs/ (accesată cel mai recent la 22 martie 2019).

32. Eun-Jae Lee, Yong-Hwan Kim, Dong-Wha Kang et al., „Deep into the brain: artificial intelligence în stroke imaging”, *Journal of Stroke*, vol. 19, nr. 3/2017, pp. 277 – 285, on-line la adresa https://www.nebi. nim. nih.gov/pâne/articles/ PMC5647643/ (accesată cel mai recent la 22 martie 2019).

33. Taylor Kubota, „Deep learning algorithm does as well as dermatologists în identifying skin cancer”, *Stanford News*, 25 ianuarie 2017, on-line la adresa https://news. stanford.edu/2017/01/25/artificial-intelligence-used-identify-skincancer/ (accesată cel mai recent la 22 martie 2019).

34. Jo Best, „IBM Watson: the inside story of how the *Jeopardywining* supercomputer was bom, and what it wants to do next”, *Tech Republic*, f.a., on-line la adresa https://www.techrepublic.com/article/ibm-watson-the-inside-storyof-how-the-jeopardy-wining-supercomputer-was-bomand-what-it-wants-to-do-next/ (accesată cel mai recent la 22 martie 2019).

35. Jennings Brown, „Why everyone is hating on IBM Watson, including the people who helped make it”, *Gizmodo*, 14 august 2017, on-line la adresa https://www.gizmodo.com. auz2017/08/why-everyone-is-hating-on-watsonincluding-the-people-who-helped-make-it/ (accesată cel mai recent la 22 martie 2019).

36. Vezi https://www.theregister.co.uk/2017/02/20/watson cancerbustingtrialonholdafterdamningaudit report/ (accesată cel mai recent la 22 martie 2019).

37. Casey Ross şi Ike Swetlitz, „IBM pitched its

Watson supercomputer as a revolution in cancer care. It's nowhere close", *STAT*, 5 septembrie 2017, on-line la adresa https://www.statnews.com/2017/09/05/watson-ibm-cancer/ (accesată cel mai recent la 22 martie 2019).

38. Tomoko Ōtake, „Big data used for rapid diagnosis of rare leukemia case în Japan", *Japan Times*, 11 august 2016, on-line la adresa https://www.japantimes.co.jp/news/2016/08/II/national/science-health/ibm-big-dataused-for-rapid-diagnosis-of-rare-leukemia-case-in-japan/. Wf8 ShOOMQ8 (accesată cel mai recent la 22 martie 2019).

39. *Amyotrophic lateral sclerosis* (ALS) (engl.), „scleroză laterală amiotrofică". Vezi „Researchers validate five new genes responsible for ALS", *Science Daily*, 1 decembrie 2017, on-line la adresa https://www.sciencedaily.com/releases/2017/12/171201104101.htm (accesată cel mai recent la 22 martie 2019).

40. John Freedman, „A reality check for IBM's AI ambitions", *MIT Technology Review*, 27 iunie 2017.

38. *Asthma facts and statistics*, Asthma UK, 2016, on-line la adresa https://www.asthma.org.uk/about/media/factsand-statistics/; *Asthma in the US*, Centers for Disease Control and Prevention, mai 2011, on-line la adresa https://www.cde.gov/vitalsigns/asthma/index.html (accesată cel mai recent la 22 martie 2019).

39. „Schoolgirl, 13, who died of asthma attack was making regular trips to A & E and running out of medication – but was NEVER referred to a specialist even when her lips turned blue, mother tells inquest", *Daily Mail*, 13 octombrie 2015, on-line la adresa http://www.dailymail.co.uk/news/article-3270728/Schoolgirl-13-died-asthma-attack-notreferred-specialist-lips-tumed-blue.html (accesată cel mai recent la 22 martie 2019).

40. My *Data*, *My care*: *How Better Use of Data Improves Health and Wellbeing*, Richmond Group of Charities, Londra, ianuarie 2017, on-line la adresa

https://Richmondgroupofeharities.org.uk/publications (accesată cel mai recent la 22 martie 2019).

41. Terence Carney, „Regulation 28: report to prevent future deaths”, raportul medico-legal în cazul Tamarei Mills, 29 octombrie 2015, on-linelaadresa https://www.judiciary.gov.uk/publications/tamara-mills/ (accesată cel mai recent la 22 martie 2019).

42. Jamie Grierson şi Alex Bem, „Doctors using Snapchat to send patient scans to ea ch other, panel finds”, *Guardian*, 5 iulie 2017, on-line la adresa https://www.theguardian.com/technology/2017/jul/05/doctors-using-snapchat-tosend-patient-scans-to-each-other-panel-finds (accesată cel mai recent la 22 martie 2019).

43. Chiar dacă reuşeşti să rezolvi toate aceste chestiuni, uneori pur şi simplu nu există informaţii. Există mii de boli rare care au la bază cauze genetice şi care sunt efectiv unice. Medicii întâmpină dificultăţi imense în descoperirea vreunei astfel de boli, pentru că nu le-au mai întâlnit niciodată. Nici toţi algoritmii din lume nu pot rezolva problema atunci când dispun de eşantioane extrem de reduse.

44. Hal Hodson, „Revealed: Google AI has access to huge haul of NHS patient data”, *New Scientist*, 29 aprilie 2016, on-line la adresa http’s: //www.newscientist.com/article/2086454 – revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/ (accesată cel mai recent la 22 martie 2019).

45. De fapt, cea mai mare parte a vinii pentru chestiunea aşa-zisei „inadecvări juridice” a fost pusă în seama Royal Free Trust, care a fost puţin cam nerăbdător să se asocieze cu cea mai cunoscută companie de inteligenţă artificială din lume. Vezi scrisoarea lui Dame Fiona Caldicott, şefa protecţiei datelor naţionale, care a ajuns la Sky News: Alex Martin, „Google received 1.6 million NHS patients data on an «inappropriate legal basis»”, Sky News, 15 mai 2017, on-line la adresa https://photos.

google.com/share/AFIQipMdd5VTKORNQIAC3Ddal52 blMGOvPD4P3 x 4 xbqmj OZA 01 rbKyxfkfypută Pvqdă/photo/AFI Qip Pl MJMXkRyy3 IuFHasil QHYEknKgnHFOFEy4T? key = U2 pZUDM4 bmo5RHhKYV – pta DlkbehfvFh4 Rmliv UVR (accesată cel mai recent la 22 martie 2019).

46. Denis Campbell, „Surgeons attack plâns to delay treatment to obese patients and smokers", *Guardian*, 29 noiembrie 2016, on-line la adresa https://www.theguardian.com/ society/2016/nov/29/surgeons-nhs-delay-treatmentobese-patients-smokers-york (accesată cel mai recent la 22 martie 2019).

47. Nir Eyal, „Denial of treatment to obese patients: the wrong policy on personal responsability for health", *International Journal of Health Policy and Management*, vol. 1, nr. 2/ august 2013, pp. 107 – 110, on-line la adresa https://www.nebi. nim. nih.gov/pâne/articles/PMC3937915/ (accesată cel mai recent la 22 martie 2019).

48. Pentru descrierea procedurilor, vezi http://galton.org/essays/1880 – 1889/galton-l 884-jaigi-anthro-lab.pdf (accesată cel mai recent la 22 martie 2019).

52. Francis Galton, „On the Anthropometric Laboratory at the late internaţional health exhibition", *Journal of the Anthropological Institute of Great Britain and Ireland*, vol. 14/1885, pp. 205 – 221.

53. „Taste", la adresa https://permalinks.23 andme.com/pdf/ samplereporttraits.pdf (accesată cel mai recent la 22 martie 2019).

54. „Sneezing on summer solstice?", *23 and McBlogf* 20 iunie 2012, on-line la adresa https://blog.23 andme.com/healthtraits/sneezing-on-summer-solstice/ (accesată cel mai recent la 22 martie 2019).

55. „Fiind out what your DNA says about your health, traits and ancestry", 23 and Me, on-line la adresa https://www.23 andme.com/en-gb/dna-health-ancestry/ (accesată cel mai recent la 22 martie 2019).

56. Kristen v. Brown, „23 and Me is selling your data but not how you think”, *Gizmodo*, 14 aprilie 2017, on-line la adresa https://gizmodo.com/23 andme-is-selling-your-databut-not-how-you-think-1794340474 (accesată cel mai recent la 22 martie 2019).

57. Michael Grothaus, „How23 and Me is monetizing your DNA”, *Fast Company*, 15 ianuarie 2015, on-line la adresa https://www.fasteompany.com/3040356/what-23 andme-is-doing-with-all-that-dna (accesată cel mai recent la 22 martie 2019).

58. Rob Stein, „Found on the Web, with DNA: a boy’s father”, *Washington Post*, 13 noiembrie 2005, on-line la adresa http://www.Washingtonpost.com/wp-dyn/content/article/2005/11/12/AR2005 111 200958.html (accesată cel mai recent la 22 martie 2019).

59. După un test ADN, tânărul a aflat că o anumită secvenţă repetitivă de pe cromozomul său Y – cel care se transmite de la tată la fiu – se mai întâlnea la două persoane purtând acelaşi nume de familie (rude îndepărtate din partea tatălui). Acel nume de familie, împreună cu locul şi data naşterii tatălui său au fost suficiente pentru a-i da acestuia de urmă.

49. M. Gymrek, A.L. McGuire, D. Golan, E. Halperin şi Y. Erlich, „Identifying personal genomes by surname inference”, *Science*, vol. 339, nr. 6117/ianuarie 2013, pp. 321 – 324, on-line la adresa https://www.nebi. nim. nih.gov/pubmed/23329047 (accesată cel mai recent la 22 martie 2019).

50. În prezent, testele genetice pentru bolii Huntington nu sunt disponibile prin intermediul kit-urilor pentru analiza ADN-ului prezente pe piaţă.

51. Matthew Herper, „23 and Me rides again: FDA clears genetic tests to predict disease risk”, *Forbes*, 6 aprilie 2017, on-line la adresa https://www.forbes.com/sites/ matthewherper/2017/04/06/23 andme-rides-again-fdaclears-genetic-tests-to-predict-disease-risk/302 aea624

fde (accesată cel mai recent la 22 martie 2019).

CAPITOLUL 5 – Automobilele autonome

1. DARPA, *Grand Challenge 2004*: *Final Report*, Defence Advanced Research Projects Agency, Arlington, 30 iulie 2004, on-line la adresa http://www.esc. whs.mil/ Portals/54/Documents/FOID/Reading%20 Room/ DARPA/15-F-0059GC2004FINALRPT7 – 30 – 2004.pdf (accesată cel mai recent la 22 martie 2019).

2. *The Worldwide Guide to Movie Locations*, *7* septembrie 2014, on-line la adresa http://www.movie-locations.com/ movies/k/Killbillvol2.html. WkYiqr TQOQ8 (accesată cel mai recent la 22 martie 2019).

3. Mariella Moon, *What you need to know about DARPA, the Pentagon's mad science division*, Engadget, 7 iulie 2014, online la adresa https://www.engadget.com/2014/07/07/darpa-explainer/ (accesată cel mai recent la 22 martie 2019).

4. DARPA, *Urban Challenge*: *Overview*, on-line la adresa http://archive. darpa.mil/grandehallenge/overview.html (accesată cel mai recent la 22 martie 2019).

5. Sebastián Thrun, „Wining the DARPA Grand Challenge, 2 August 2006", postat pe YouTube la 8 octombrie 2007, on-line la adresa https://www.youtube.com/ watch?v=j8 zj5 lBpFTY (accesată cel mai recent la 22 martie 2019).

6. DARPA, Urban Challenge: Overview, ed. cit.

7. „DARPA Grand Challenge 2004 – road to…", prezentare postată pe YouTube la 22 ianuarie 2014, on-line la adresa https://www. youtube.com/watch?v=FabJ5 sPPmcl (accesată cel mai recent la 22 martie 2019).

8. Alex Davies, „An oral history of the DARPA Grand Challenge, the grueling robot race that launched the self-driving car", *Wired*, 8 martie 2017, on-line la adresa https://www.wired.com/story/darpa-grand-challenge-2004-oralhistory/ (accesată cel mai recent la 22 martie 2019).

9. „Desert race too tough for robots", ştire BBC News

din 15 martie, 2004, on-line la adresa http://news. bbe.co. Uk/l/ hi/technology/3512270. stm (accesată cel mai recent la 22 martie 2019).

10. Alex Davies, op. cit.

11. Denise Chow, „DARPA and drone care: how the US military spawned self-driving car revolution", *Live Science*, 21 martie 2014, on-line la adresa https://www.livescience.com/44272-darpa-self-driving-car-revolution.html (accesată cel mai recent la 22 martie 2019).

12. Joseph Hooper, „From Darpa Grand Challenge 2004 DARPA's debacle in the desert", *Popular Science*, 4 iunie 2004, on-line la adresa https://www.popsci.com/scitech/ article/2004 – 06/darpa-grand-challenge-2004 darpas-debacle-desert (accesată cel mai recent la 22 martie 2019).

13. Alex Davies, op. cit.

14. DARPA, *Report to Congress*: *DARPA Prize Authority*. *Fiscal Year 2005 Report in Accordance with 10 U.S.C. 2374 a*, martie 2006, on-line la adresa http://archive. darpa.mil/ grandehallenge/docs/grandehallenge2005 reportto congress.pdf (accesată cel mai recent la 22 martie 2019).

15. Alan Ohnsman, „Bosch and Daimler to partner to get driverless taxis to market by early 2020 s", *Forbes*, 4 aprilie 2017, on-line la adresa https://www.forbes.com/ sites/alanohnsman/2017/04/04/bosch-and-daimlerpartner-to-get-driverless-taxis-to-market-by-early2020 s/3 obec7 e63 c4 b (accesată cel mai recent la 22 martie 2019).

16. Ford, *Looking Further*: *Ford Will Have a Fully Autonomous Vehicle în Operation by 2021*, on-line la adresa https://corporate. Ford.com/inovation/autonomous-2021.html (accesată cel mai recent la 22 martie 2019).

17. John Markoff, „Should your driverless car hit a pedestrian to save your life?", *New York Times*, 23 iunie 2016, on-line la addresa http's: //www.nytimes.com/2016/06/24/technology/should-your-

driverless-car-hit-a-pedestrian-tosave-your-life.html (accesată cel mai recent la 22 martie 2019).

18. Clive Thompson, Anna Wiener, Ferris Jabr, Rahawa Haile, Geoff Manaugh, Jamie Lauren Keiles, Jennifer Kahn şi Malia Wollan, „Full tilt: when 100 per cent of care are autonomous", *New York Times*, 8 nov. 2017, https://www.nytimes.com/interactive/2017/11/08/magazine /tech-design-autonomous-future-cars-l OO-percent-augmented-reality-policing. htmithe-end-of-roadkill (accesată cel mai recent la 22 martie 2019).

19. Peter Campbell, „Trucks headed for a driverless future: unions warn that millions of drivers jobs will be disrupted", *Financial Times*, 31 ian. 2018, on-line la adresa https://www.fi.com/content/7686 ea Ae-eOdd-lle7-aOd4 – 0944 c5 f49 e46 (accesată cel mai recent la 22 martie 2019).

20. Markus Maurer, J. Christian Gerdes, Barbara Lenz şi Hermann Winner, *Autonomous Driving*: *Tehnical*, *Legal and Social Aspects*, Springer, New York, mai 2016, p 48.

21. Stephen Zavestoski şi Julian Agyeman, *Incomplete Streets: Processes*, *Practices*, *and Possibilities*, Routledge, Londra, 2015, p. 29.

22. Markus Maurer, J. Christian Gerdes, Barbara Lenz şi Hermann Winner, op. cit. p. 53.

23. David Rooney, *Self-guided Cars*, Muzeul Ştiinţei, Lonfta, 27 aug. 2009, on-line la adresa https://blog. sciencemuseum.org.uk/self-guided-cars/ (accesată cel mai recent la 22 martie 2019).

24. Blake Z. Rong, „How Mercedes sees into the future", *Autoweek*, 22 ianuarie 2014, pn-line la adresa http://autoweek.com/article/carnews/how-mercedes-sees-future (accesată cel mai recent la 22 martie 2019).

25. Dean A. Pomerleau, *ALVINN*: *An Autonomous Land Vehicle* in a *Neural Network*, CMU-CS-89 – 107 (Pittsburgh: Carnegie Mellon University, ian. 1989), http://repository. cmu.edu/cgi/viewcontent. cgi? article =

2874 & context = compsci.

26. Joshua Davis, „Say hello to Stanley”, *Wired*, 1 ianuarie 2006, on-line la adresa https://www.wired.com/2006/01/ Stanley/ (accesată cel mai recent la 22 martie 2019); iar pentru mai multe detalii, Dean A. Pomerleau, *Neural Network Perception for Mobile Robot Guidance*, Springer, New York, 2012, p. 52.

27. A. Fugueira, H. González-Jorge, S. Lagiiela, L. Diaz-Vilarino şi P. Arias, „Quantifying the influence of rain în LIDAR performance”, *Measurement*, vol. 95/ianuarie 2017, pp. 143 – 148, DOI: https://doi. Org/10.1016/j. mea* surement.2016.10.009; https://www.sciencedirect.com/ science/article/pii/S0263224116305577 (accesată cel mai recent la 22 martie 2019).

28. Chris Williams, „Stop lights, sunsets, junctions are tough work for Google’s robo-cars”, *The Register*, 24 august 2016, on-line la adresa https://www.theregister. co.uk/2016/08/24/googleselfdrivingcarproblems/ (accesată cel mai recent la 22 martie 2019).

29. Novatel, *IMU Errors and Their Effects*, on-line la adresa https://www.novatel.com/assets/Documents/Bulletins/ APN064.pdf (accesată cel mai recent la 22 martie 2019).

30. Teorema în sine nu este decât o ecuaţie care face legătura între probabilitatea unei *ipoteze* dată fiind observarea unor indicii, şi probabilitatea existenţei acelor indicii dată fiind ipoteza. O expunere introductivă mai cuprinzătoare poate fi găsită on-line la adresa https://arbitrar.com/p/ bayesrule/? l = lzq (accesată cel mai recent la 22 martie 2019).

31. Sharon Bertsch McGrayne, The Theory That Would Not Die: How Bayes’ Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy, Yale University Press, New Haven, 2011.

32. M. Bayes şi M. Price, *An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr.*

Bayes, F.R.S. Communicated by Mr. Price, in a *Letter to John Canton, A.M.F.R.S.* (1763), copie digitală încărcată în archive.org la 2 august 2011,https://archive.org/details/philtranş09948070 (accesată cel mai recent la 22 martie 2019).

33. Michael Taylor, „Self-driving Mercedes-Benzes will prioritize occupant safety over pedestrians", *Car and Driver*, 7 octombrie 2016, on-line la adresa https://blog.caranddriver.com/self-driving-mercedes-will-prioritize-occupant-safety-over-pedestrians/ (accesată cel mai recent la 22 martie 2019).

34. Jason Kottke, *Mercedes' Solution to the Trolley Problem*, Kottke.org, 24 octombrie 2016, on-line la adresa https://kottke.org/16/10/mercedes-solution-to-the-trolley-problem (accesată cel mai recent la 22 martie 2019).

35. Jean-François Bonnefon, Azim Shariff şi Iyad Rahwan (2016), „The social dilema of autonomous vehicles", *Science*, vol. 35/24 iunie 2016, on-line la adresa DOI 10.1126/science. aaf2654; http's: //arxiv.org/pdf/1510.03346.pdf (accesată cel mai recent la 22 martie 2019).

5. Toate citatele provin dintr-o discuţie particulară purtată cu Paul Newman.

6. Naaman Zhou, „Volvo admits its self-driving care are confused by kangaroos", *Guardian*, 1 iulie 2017, on-line la adresa https://www.theguardian.com/technology/2017/jul/01/volvo-admits-its-self-driving-cars-are-confused-bykangaroos (accesată cel mai recent la 22 martie 2019).

7. Toate citatele provin dintr-o discuţie particulară purtată cu Jack Stil goe.

8. Jeff Sabatini, „The one simple reason nobody is talking realistically about driverless care", *Car and Driver*, octombrie 2017, http's: //www.caranddriver.com/features/the-onereason-nobody-is-talking-realistically-aboutdriverless-cars-feature (accesată cel mai recent la 22 martie 2019).

9. William Langewiesche, „The human factor”, *Vanity Fair*, 17 septembrie 2014, on-line la adresa https://www.vanityfair.com/news/business/2014/10/air-france-flight447-crash (accesată cel mai recent la 22 martie 2019).

10. Bureau d’Enquêtes et d’Analyses pour la Securitie de l’Aviation Civile, *Final Report on the Accident on îşi June 2009 to the Airbus A330 – 203 registered F-GZCP operated by Air France Flight AF447 Rio de Janeiro – Paris*, ediţia în limba engleză, Paris, actualizată în iulie 2012, on-line la adresa https://www.bea.aero/docspa/2009/f-cp090601. en/pdf/f-cp090601.en.pdf (accesată cel mai recent la 22 martie 2019).

11. Ibidem.

12. William Langewiesche, op. cit.

13. Ibidem.

14. Jeff Wise, „What really happened aboard Air France 447”, *Popular Mechanics*, *6* decembrie 2011, on-line la adresa http://www.popularmechanics.com/flight/a3115/whatreally-happened-aboard-air-ffance-447 – 6611877/ (accesată cel mai recent la 22 martie 2019).

15. William Langewiesche, op. cit.

47. Jeff Wise, op. cit.

48. Lisanne Bainbridge, „Ironies of automation”, *Automatica*, vol. 19, nr. 6/noiembrie 1983, pp. 775 – 9, on-line la adresa https://www.sciencedirect.com/science/article/pii/0005109883900468 (accesată cel mai recent la 22 martie 2019).

49. Ibidem.

50. Alex Davies, „Everyone wants a level 5 self-driving car – here’s what that means”, *Wired*, 26 iulie 2016.

51. Justin Hughes, „Car autonomy levels explained”, *The Drive*, 3 noiembrie 2017, on-line la adresa

http://www.thedrive.com/sheetmetal/15724/what-are-these-levels-ofautonomy-anyway (accesată cel mai recent la 22 martie 2019).

52. Lisanne Bainbridge, op. cit.

53. Jack Stil goe, „Machine learning, social learning and the governance of self-driving care", *Social Studies of Science*, vol. 48, nr. 1/2017, pp. 25 – 56.

54. Eric Tingwall, „Where are autonomous care right now? Four sistems tested", *Car and Driver*, octombrie 2017, on-line la adresa http's: //www.caranddriver.com/features/where-are-autonomous-cars-right-now-four-systemstested-feature (accesată cel mai recent la 22 martie 2019).

55. Tracey Lindeman, „Using an orange to fool Tesla's autopilot is probably a really bad ideea", *Motherboard*, 16 ianuarie 2018, on-line la adresa https://motherboard. vice.com/enus/article/a Ana9 p/tesla-autosteer-orange-hack (accesată cel mai recent la 22 martie 2019).

56. Daisuke Wakabayashi, „Über's self-driving care were struggling before Arizona Crash", *New York Times*, 23 martie 2018, on-line la adresa https://www.nytimes.com/2018/03/23/technology/über-self-driving-carsarizona.html (accesată cel mai recent la 22 martie 2019).

57. Sam Levin, „Video released of Über self-driving crash that killed woman în Arizona", *Guardian*, 22 martie 2018, on-line la adresa http's: //www.theguardian.com/

technology/2018/mar/22/video-released-of-uber-selfdriving-crash-that-killed-woman-in-arizona (accesată cel mai recent la 22 martie 2019).

58. Audi, *The Audi vision of autonomous driving*, Audi Newsroom, 11 sept. 2017, on-line la adresa https://media. audiusa.com/en-us/releases/184 (accesată cel mai recent la 22 martie 2019).

59. P. Morgan, C. Alford şi G. Parkhurst, *Handover Issues* în *Autonomous Driving*: *A Literature Review*.

Project Report, University of the West of England, Bristol, iunie 2016, on-line la adresa http://eprints. uwe.ac. Uk/29167/l/ VenturerWP5.2 Lit%20 Review Handover.pdf (accesată cel mai recent la 22 martie 2019).

60. William Langewiesche, op. cit.

61. Evan Ackerman, „Toyota's Gill Pratt on self-driving care and the reality of full autonomy", *IEEE Spectrum*, *23* ianuarie 2017, on-line la adresa https://spectrum. ieee.org/ cars-that-think/transportation/self-driving/toyota-gilipratt-on-the-reality-of-full-autonomy (accesată cel mai recent la 22 martie 2019).

62. Julia Pyper, „Self-driving care could cut greenhouse gas pollution", *Scientific American*, 15 septembrie 2014, on-line la adresa http's: //www.scientificamerican.com/ article/self-driving-cars-could-cut-greenhouse-gaspollution/ (accesată cel mai recent la 22 martie 2019).

63. Raphael E. Stem et al., „Dissipation of stop-and-go waves via control of autonomous vehicles: field experiments", *arXiv: 1705.01693 vi*, 4 mai 2017, on-line la adresa https://arxiv.org/abs/1705.01693 (accesată cel mai recent la 22 martie 2019).

64. Some] oe7777, „Tesla Model S forward collision warning saves the day", postare încărcată pe YouTube la 19 octombrie 2016, on-line la adresa https://www.youtube.com/watch?v=SnRp56 XivM (accesată cel mai recent la 22 martie 2019).

58. Jordan Golson şi Dieter Bohn, „All new Tesla care now have hardware for „full self-driving capabilities": but some safety features will be disabled initially", *The Verge*, 19 octombrie 2016, on-line la adresa https://www.theverge.com/2016/10/19/13340938/tesla-autopilot-updatemodel-3-elon-musk-update (accesată cel mai recent la 22 martie 2019).

59. Fred Lambert, „Tesla introduces first phase of „Enhanced Autopilot": „measured and cautious for next

several hundred million miles" – release notes", *Electrek*, 1 ianuarie 2017, on-line la adresa http's: //electrek.co/2017/01/01/ tesla-enhanced-autopilot-release-notes/ (accesată cel mai recent la 22 martie 2019).

60. DPC Cars, „Toyota Guardian and Chauffeur autonomous vehicle platform", postare încărcată pe YouTube la 27 septembrie 2017, on-line la adresa https://www.youtube.com/watch?v=IMdce KGJ90 c (accesată cel mai recent la 22 martie 2019).

61. Brian Milligan, „The most significant development since the safety belt", ştire BBC News din 15 aprilie 2018, on-line la adresa http://www.bbe.co.uk/news/business43752226 (accesată cel mai recent la 22 martie 2019).

CAPITOLUL 6 – Infracţionalitatea

1. Bob Taylor, *Crimebuster*: *Inside the Minds of Britain's Most Evil Criminals*, Piatkus, Londra, 2002, capitolul 9, „A day out from jail".

2. Ibidem.

3. Nick Davies, „Dangerous, în prison – but free to rape", *Guardian*, 5 octombrie 1999, on-line la adresa https://www.theguardian.com/uk/1999/oct/05/nickdaviest (accesată cel mai recent la 22 martie 2019).

4. João Medeiros, „How geographic profiling helps fiind serial criminals", *Wired*, noiembrie 2014, on-line la adresa http://www.wired.co.uk/article/mapping-murder (accesată cel mai recent la 22 martie 2019).

5. Nicole H. Rafter, ed., *The Origins of Criminology*: *A Reader*, Routledge, Abingdon, 2009, p. 271.

6. Luke Dormehl, The Formula: How Algorithms Solve All Our Problems… and Create More, W.H. Allen, Londra, 2014, p. 117.

7. *Ibidem*, p. 116.

8. D. Kim Rossmo, „Geographic profiling", în Gerben Bruinsma şi David Weisburd (editori), *Encyclopedia of Criminology and Criminal Justice*, Springer, New York, 2014, on-line la adresa https://link.

springer.com/referencework entry/10.1007%2F978 – 1 – 4614 – 5690 – 2678.

9. Ibidem.

10. João Medeiros, *op. tit.*

11. Ibidem.

12. „Sadic" serial rapist sentenced to eight life terms", *Independent* (Irlanda), 6 octombrie 1999, on-line la adresa http://www.independent.ie/world-news/sadistic-serialrapist-sentenced-to-eight-life-terms-26134260.html (accesată cel mai recent la 22 martie 2019).

13. Ibidem.

14. Steven C. Le Comber, D. Kim Rossmo, Ali N. Hassan, Douglas O. Fuller şi John C. Beier, „Geographic profiling as a novel spaţial tool for targeting infectious disease control", *International Journal of Health Geographies*, vol. 10, nr. 1/2011, p. 35, on-line la adresa https://www.nebi. nim. nih.gov/pâne/articles/PMC3123167/ (accesată cel mai recent la 25 martie 2019).

15. Michelle V. Hauge, Mark D. Stevenson, D. Kim Rossmo şi Steven C. Le Comber, „Tagging Banksy: using geographic profiling to investigate a modern art mystery", *Journal of Spaţial Science*, vol. 61, nr. 1/2016, pp. 185 – 90, on-line la adresa http://www.tandfonline.com/doi/abs/10.1080/ 14498596.2016.1138246 (accesată cel mai recent la 25 martie 2019).

16. Raymond Dussault, „Jack Maple: betting on intelligence", *Government Technology*, 31 martie 1999, on-line la adresa http://www.govtech.com/featured/Jack-Maple-Bettingon-Intelligence.html (accesată cel mai recent la 25 martie 2019).

1. Ibidem.

2. Ibid.

3. Nicole Gelinas, „How Bratton's NYPD saved the subway sistem", *New York Post*, 6 august 2016, on-line la adresa http://nypost.com/2016/08/06/how-brattons-nypd-saved-the-subway-system/ (accesată cel mai recent la 25

martie 2019).

4. Raymond Dussault, op. cit.

5. Andrew Guthrie Ferguson, „Predictive policing and reasonable suspición”, *Emory Law Journal*, vol. 62, nr. 2/2012, p. 259, on-line la adresa http://law. emory.edu/el/content/ volume-62/issue-2/articles/predicting-policing-andreasonable-suspición.html (accesată cel mai recent la 25 martie 2019).

6. Lawrence W. Sherman, Patrick R. Gratin şi Michael E. Buerger, „Hot spots of predatory crime: routine activities and the criminology of place”, *Criminology*, vol. 27, nr. 1/1989, pp. 27 – 56, on-line la adresa http://onlinelibrary. wiley.com/doi/10.1111 /). 1745 – 9125.1989. th00862. x/abstract (accesată cel mai recent la 25 martie 2019).

7. Toby Davies şi Shane D. Johnson, „Examining the relationship between road structure and burglary risk via quantitative network analysis”, *Journal of Quantitative Criminology*, vol. 31, nr. 3/2015, pp. 481 – 507, on-line la adresa http://discovery. uci.ac. Uk/1456293/5/Johnsonart %253A 10.1007%252 Fsl0940 – 014 – 9235 – 4.pdf (accesată cel mai recent la 25 martie 2019).

8. Michael J. Frith, Shane D. Johnson şi Hannah M. Fry, „Role of the street network în burglars spaţial decisionmaking”, *Criminology*, vol. 55, nr. 2/2017, pp. 344 – 76, online la adresa http://onlinelibrary. wiley.com/doi/10. 1 111/ 1745 – 9125.12133/full (accesată cel mai recent la 25 martie 2019).

25. Spencer Chainey, *Predictive Mapping* (*Predictive Policing*), JDI Brief, Jill Dando Institute of Security and Crime Science, University College Londra, 2012, on-line la adresa http://discovery. uci.ac.uk/1344080/3/JDIB – nefs Predictive Mappingălhainey April2012.pdf (accesată cel mai recent la 25 martie 2019).

26. Ibidem.

27. Un mic avertisment. Algoritmul Pred Pol în sine nu este accesibil publicului. Experimentul la care ne

referim în carte a fost realizat de aceiaşi matematicieni care au creat Pred Pol, cu ajutorul unei tehnici care se potriveşte cu modul în care este descris software-ul brevetat. Toate indiciile sugerează că amândoi algoritmii sunt la fel, dar, strict vorbind, nu putem fi *absolut* siguri.

28. G.O. Mohler, M.B. Short, Sean Malinowski, Mark Johnson, G.E. Tita, Andrea L. Bertozzi şi P.J. Brantingham, „Randomized controlled field trials of predictive policing", *Journal of the American Statistical Association*, vol. 110, nr. 512/2015, pp. 1399 – 1411, on-line la adresa http://www.tandfonline.com/doi/abs/10.1080/01621459.2015.10777 10 (accesată cel mai recent la 25 martie 2019).

29. Departamentul de analiză al serviciului administrativ al Poliţiei din Kent, *Pred Pol Operaţional Review*, 2014, on-line la adresa http://www.statewatch.org/docbin/uk2014-kent-police-predpol-op-review.pdf (accesată cel mai recent la 25 martie 2019).

30. G.O. Mohler, M.B. Short, Sean Malinowski, Mark Johnson, G.E. Tita, Andrea L. Bertozzi şi P.J. Brantingham, op. cit.

31. Departamentul de analiză al serviciului administrativ al Poliţiei din Kent, Marea Britanie, *Pred Pol Operaţional Review*: Iniţial *Findings*, 2013, on-line la adresa https://www.whatdotheyknow.com/request/181341/response/454199/attach/3/13%2010%20888%20Appendix, pdf (accesată cel mai recent la 25 martie 2019).

32. Ibidem.

9. Acesta nu era Pred Pol, ci un algoritm mai simplu care utiliza tot ideea efectelor de „semnal" şi „intensificare". Vezi Matthew Fielding şi Vincent Jones, „Disrupting the optimal forager: predictive risk mapping and domestic burglary reduction in Trafford, Greater Manchester", *International Journal of Police Science and Management*, vol. 14, nr. 1/2012, pp. 30 – 41.

10. Joe Newbold, „Predictive policing", „preventative policing" or „intelligence led policing". What is the

future?”, Proiect consultativ prezentat spre evaluare de către programul de MBA Warwick, Warwick Business School, 2015.

11. Informaţii din 2016: COMPSTAT, *Citywide Profile 12/04/16 – 12/31/16*, on-line la adresa http://assets. lapdonline.org/assets/pdf/123116 cityprof.pdf (accesată cel mai recent la 25 martie 2019).

12. Ronald V. Clarke şi Mike Hough, *Crime and Police Efectiveness*, Home Office Research Study nr. 79, HMSO, Londra, 1984, on-line la adresa https://archive.org/stream/ opl276605 – 1001/opl276605 – 1001 divu.txt (accesată cel mai recent la 25 martie 2019), apud Tom Gash, *Criminal*: *The Truth about Why People Do Bad Things*, Allen Lane, Londra, 2016.

13. Departamentul de analiză al serviciului administrativ al Poliţiei din Kent, op. cit.

14. Pred Pol, „Recent examples of crime reduction”, 2017, online la adresa http://www.predpol.com/results/ (accesată cel mai recent la 25 martie 2019).

15. Aaron Shapiro, „Reform predictive policing, *Nature*, vol. 541, nr. 7638/25 ianuarie 2017, on-line la adresa http://www.nature.com/news/reform-predictive-policing-l.21338 (accesată cel mai recent la 25 martie 2019).

16. Chicago Dată Portal, Strategic Subiect List, on-line la adresa http’s: //data. cityofehicago.org/Public-Safety/Strategic-Subject-List/4 aki-rănp (accesată cel mai recent la 25 martie 2019).

41. Jessica Saunders, Priscilla Hunt şi John Hollywood, „Predictions put into practice: a quasi-experimental evaluation of Chicago’s predictive policing pilot”, *Journal of Experimental Criminology*, vol. 12, nr. 3/2016, pp. 347 – 371.

42. Copblock, „Innocent man arrested for robbery and assault, spends two months în Denver jail”, 28 aprilie 2015, on-line la adresa https://www.copblock.org/122644/man-arrested-for-

robbery-assault-he-didnt-commit-spends-twomonths-indenver-jail/ (accesată cel mai recent la 25 martie 2019).

43. Ibidem.

44. Ava Kofman, „How a facial recognition mismatch can ruin your life", *The Intercept*, 13 octombrie 2016.

45. Ibidem.

46. Copblock, „Denver police, „Don't fck with the biggest gang în Denver" before beating man wrongfully arrested – TWICE!!", 30 ianuarie 2016, on-line la adresa https://www.copblock.org/152823/Denver-police-fck-up-again/ (accesată cel mai recent la 25 martie 2019).

47. De fapt, povestea lui Talley este chiar mai răscolitoare decât am prezentat-o aici în rezumat. După ce a petrecut două luni în închisoare, când a fost arestat iniţial, Talley a fost eliberat fără să i se aducă vreo acuzaţie. Un an mai târziu – în momentul în care el trăia deja într-un adăpost – a fost arestat pentru a doua oară. De această dată acuzaţiile nu au fost retrase, iar FBI a depus mărturie împotriva sa. Procesul a încetat atunci când funcţionara de la bancă şi-a dat seama că lui Talley îi lipseau aluniţele pe care le observase pe mâinile jefuitorului când acesta le întinsese peste ghişeu, şi a declarat în instanţă: „Acesta nu este individul care m-a jefuit". Acum Talley cere despăgubiri de zece milioane de dolari. Pentru relatarea completă, vezi Ava Kofman, op. cit.

48. Justin Huggler, „Facial recognition software to câtch terrorists being tested at Berlin station", *Telegraph*, 2 august 2017, on-line la adresa http://www.telegraph.co.uk/

news/2017/08/02/facial-recognition-software-catchterrorists-tested-berlin-station/ (accesată cel mai recent la 25 martie 2019).

17. David Kravets, „Driver's license facial recognition tech leads to 4.000 New York arrests", *Ars Tehnica*, 22 august 2017, on-line la adresa https://arstehnica.com/tech-policy/2017/08/biometrics-leads-to-thousands-of-a-ny-arrests-for-fraud-identity-theft/ (accesată cel mai recent la

25 martie 2019).

18. Ruth Mosalski, „The first arrest using facial recognition software has been made”, *Wales Online*, 2 iunie 2017, on-line la adresa http://www.walesonline.co.uk/news/ local-news/first-arrest-using-facial-recognition-13126934 (accesată cel mai recent la 25 martie 2019).

19. Sebastián Anthony, „UK police arrest man via automatic face recognition tech”, *Ars Tehnica*, 6 iunie 2017, on-line la adresa https://arstehnica.com/tech-policy/2017/06/ police-automatic-face-recognition (accesată cel mai recent la 25 martie 2019).

20. David White, Richard I. Kemp, Rob Jenkins, Michael Matheson şi A. Mike Burton, „Passport officers errors în face matching”, *PLOSOne*, 18 august 2014, on-line la adresa http://joumals. plos.org/plosone/article? id = 10.1371 /journal. pone.0103510 se (accesată cel mai recent la 25 martie 2019).

21. Teghan Lucas şi Maciej Henneberg, „Are human faces unique? A metric approach to finding single individuals without duplicates în large samples”, *Forensic Science International*, vol. 257/decembrie 2015, pp. 514 el-514. e6, on-line la adresa http://www.sciencedirect.com/science/ article/pii/S0379073815003758 (accesată cel mai recent la 25 martie 2019).

22. Zaria Gorvett, „You are surprisingly likely to have a living doppelganger”, BBC Future, 13 iulie 2016, on-line la adresa http://www.bbe.com/future/story/20160712-youare-surprisingly-likely-to-have-a-living-doppelganger (accesată cel mai recent la 25 martie 2019).

55. „Eyewitness misidentification”, The Innocence Project, on-line la adresa https://www.innocenceproject.org/causes/eyewitness-misidentification (accesată cel mai recent la 25 martie 2019).

56. Douglas Starr, „Forensics gone wrong: when DNA

snares the innocent", *Science*, 7 martie 2016, on-line la adresa http://www.sciencemag.org/news/2016/03/forensicsgone-wrong-when-dna-snares-innocent (accesată cel mai recent la 25 martie 2019).

57. Asta nu înseamnă că, dacă se apelează la ADN, falsele identificări sunt excluse – ele se mai întâmplă; înseamnă numai că ai o armă puternică de partea ta pentru a le face să fie cât mai rare cu putinţă.

58. Richard W. Vorder Bruegge, *Individualization of People from Images*, Divizia de Tehnologie Operaţională a FBI, Unitatea de analiză criminalistică audio, video şi imagine, Quantico, 12 decembrie 2016, on-line la adresa https://www.nist.gov/sites/default/files/documents/2016/12/12/ vorderbruegge-face.pdf (accesată cel mai recent la 25 martie 2019).

59. Lance Ulanoff, „The iPhone X can't tell the difference between twins", *Maskable UK*, 31 octombrie 2017, online la adresa http://mashable.com/2017/10/31/puttingiphone-x-face-id-to-twin-test/A87 kA26 aAqqQ (accesată cel mai recent la 25 martie 2019).

60. Kif Leswing, „Apple says the iPhone X's facial recognition sistem isn't for kids", *Business Insider UK*, 27 septembrie 2017, on-line la adresa http://uk. businessinsider.com/ apple-says-the-iphone-xs-face-id-is-less-accurate-onkids-under-13 – 2017 – 9 (accesată cel mai recent la 25 martie 2019).

61. Andy Greenberg, „Watch a 10-year-old's face unlock his mom's iPhone X", *Wired*, 14 noiembrie 2017, on-line la adresa https://www.wired.com/

story/10-year-old-face-id-unlocks-mothers-iphone-x/ (accesată cel mai recent la 25 martie 2019).

23. „Bkav's new mask beats Face ID în „twin way": severity level caised, do not use Face ID în business transactions", Bkav Corporation, 27 noiembrie 2017, on-line la adresa http://www.bkav.com/dt/top-

news/-/viewcontent/ content/103968/bkav%EL%BI%BDs-new-mask-beatsface-id-in-twin-way-severity-level-raised-do-not-use-faceid-in-business-transactions (accesată cel mai recent la 25 martie 2019).

24. Mahmood Sharif, Śruti Bhagavatula, Lujo Bauer şi Michael Reiter, „Accesorize to a crime: real and stealthy attacks on state-of-the-art face recognition”, lucrare prezentată la Conferinţa ACM SIGSAC, 2016, on-line la adresa https://www.ce. cmu.edu/-sbhagava/papers/face-rec-ccsl6.pdf (accesată cel mai recent la 25 martie 2019).

25. Ira Kemelmacher-Shlizerman, Steven M. Seitz, Daniel Miller şi Evan Brossard, *The Mega Face Benchmark*: *1 Million Faces for Recognition at Scale*, Computer Vision Foundation, 2015, on-line la adresa https://arxiv.org/abs/1512.00596 (accesată cel mai recent la 25 martie 2019).

26. „Half of all American adults are în a police face recognition database, new report finds”, comunicat de presă, articol publicat de Facultatea de Drept a Universităţii din Georgetown, 18 octombrie 2016, https://www.law. Georgetown.edu/news/press-releases/half-of-all-american-adults-arein-a-police-face-recognition-database-new-report-finds. din (accesată cel mai recent la 25 martie 2019).

27. Josh Chin şi Liza Lin, „China’s all-seeing surveillance state is reading its citizens faces”, *Wall Street Journal* 6 iunie 2017, on-line la adresa https://www.wsi.com/articles/ the-all-seeing-surveillance-state-feared-in-the-west-is-areality-in-china-1498493020 (accesată cel mai recent la 25 martie 2019).

67. Daniel Miller, Evan Brossard, Steven M. Seitz şi Ira Kemelmacher-Shlizerman, *The Mega Face Benchmark*: *1 Million Faces for Recognition at Scale*, 2015, on-line la adresa https://arxiv.org/pdf/1505.02108.pdf (accesată cel mai recent la 25 martie 2019).

68. Ibidem.

69. Mega Face and MF2: Million-Scale Face

Recognition, „Most recent public results”, 12 martie 2017, http://megaface. ce. Washington.edu/; „Leading facial recognition platform Tencent Youâu Lab smashes records în Mega Face facial recognition challenge”, Cision PR Newswire, 14 aprilie 2017, on-line la adresa http://www.pânewswire.com/news-releases/leading-facial-recognition-platform-tencentyoutu-lab-smashes-records-in-megaface-facial-recognition-challenge-300439812.html (accesată cel mai recent la 25 martie 2019).

70. Dan Robson, „Facial recognition a sistem problem gamblers can’t beat?”, *The Star.com*, 12 ianuarie 2011, https://www.thestar.com/news/gta/2011/01/12/ facialrecognitionasystemproblemgamblerscant beat.html (accesată cel mai recent la 25 martie 2019).

71. British Retail Consortium, *2016 Retail Crime Survey*, BRE, Londra, februarie 2017, on-line la adresa https://bre.org.uk/ media/116348/10081 – bre-retail-crime-survey-2016 all-graphics-latest.pdf (accesată cel mai recent la 25 martie 2019).

72. D & D Daily, *The D & D Daily’s 2016 Retail Violent Death Report*, 9 martie 2017, on-line la adresa http://www.d-ddaily.com/archivesdaily/Daily Special Report03 – 09-17F.htm (accesată cel mai recent la 25 martie 2019).

73. Joan Gumey, „Walmart’s use of facial recognition tech to spot shoplifters caises privacy concems”, iQ Metrix, 9 noiembrie 2015, on-line la adresa http://www.iqmetrix.com/ blog/walmarts-use-of-facial-recognition-tech-to-spot-shoplifters-raises-privacy-concems (accesată cel mai recent la 25 martie 2019).

74. Andy Coghlan şi James Randerson, „How far should fingerprints be trusted?”, *New Scientist*, 14 sept. 2005, on-line la adresa https://www.newscientist.com/article/dn8011 how-far-should-fingerprints-be-trusted/ (accesată cel mai recent la 25 martie 2019).

75. Phil Locke, „Blood spatter – evidence?”, *The*

Wrongful Convictions Blog, 30 aprilie 2012, on-line la adresa https://wrongfulconvictions blog.org/2012/04/30/blood-spatterevidence/ (accesată cel mai recent la 25 martie 2019).

76. Michael Shermer, „Can we trust crime forensics?", *Scientific American*, 1 septembrie 2015, on-line la adresa https://www. scientificamerican.com/article/canwe-trust-crimeforensics/ (accesată cel mai recent la 25 martie 2019).

77. Consiliul Naţional de Cercetare al Academiei Naţionale de Ştiinţe, *Strengthening Forensic Science in the United States*: *A Path Forward*, National Academies Press, Washington, DC, 2009, p. 7, on-line la adresa https://www.nqrs.gov/pdffilest/nu/grants/228091.pdf (accesată cel mai recent la 25 martie 2019).

78. Colin Moynihan, „Hammer attacker sentenced to 22 years în prison", *New York Times*, 19 iulie 2017, on-line la adresa https://www.nytimes.com/2017/07/19/nyregion/hammer-attacker-sentenced-to-22-years-in-prison.html? meubz = O (accesată cel mai recent la 25 martie 2019).

79. Jeremy Tanner, „David Baril charged în hammer attacks after police-involved shooting", *Pixil*, 14 mai 2015, on-line la adresa http://pixil.com/2015/05/14/david-baril-chargedin-hammer-attacks-after-police-involved-shooting/ (accesată cel mai recent la 25 martie 2019).

80. „Longtime fugitive captured juggler was on the run for 14 years", informaţii FBI, 12 august 2014, on-line la adresa https://www.fbi.gov/news/stories/long-time-fugitiveneil-stammer-captured (accesată cel mai recent la 25 martie 2019).

81. Pei-Sze Cheng, „I-Team: use of facial recognition technology expands as some question whether rules are keeping up", ştire NBC prezentată în emisiunea *4 NewYork* din 23 iunie 2015, on-line la adresa http://www.nbenewyork.com/news/local/Facial-Recognition – NYPD-Technology-Video-Camera-Police-Arrest-

Surveillance-309359581.html (accesată cel mai recent la 25 martie 2019).

28. Nate Berg, „Predicting crime, LAPD-style”, *Guardian*, *25* iunie 2014, on-line la adresa https://www.theguardian.com/cities/2014/jun/25/predicting-crime-lapd-los-angeles-police-data-analysis-algorithm-minority-report (accesată cel mai recent la 25 martie 2019).

CAPITOLUL 7 – *Arta*

1. Matthew J. Salganik, Peter Sheridan Dodds şi Duncan J. Watts, „Experimental study of inequality and impredictability în an artificial cultural market”, *Science*, vol. 311/10 februarie 2006, p. 854, DOI [Digital Obiect Identifier]: 10.1126/science. 1 121066, on-line la adresa https://www.princeton.edu/-mjsă/salganikdoddswattsobfull. pdf (accesată cel mai recent la 28 martie 2019).

2. Vezi site-ul dedicat Musiclab, on-line la adresa http://www. princeton.edu/-mjsă/musiclab. shtml (accesată cel mai recent la 28 martie 2019).

3. Kurt Kleiner, „Your taste în music is shaped by the crowd”, *New Scientist*, 9 februarie 2006, on-line la adresa https://www. newscientist.com/article/dn8702-your-taste-in-music-is-shaped-by-the-crowd/ (accesată cel mai recent la 28 martie 2019).

4. Bjørn Carey, „The science of hit songs”, *Live Science*, 9 februarie 2006, on-line la adresa https://www.livescience.com/7016-science-hit-songs.html (accesată cel mai recent la 28 martie 2019).

5. „Vanilla, indeed”, *True Music Facts Wednesday Blogspot*, 23 iulie 2014, on-line la adresa http://truemusicfactswednesday.blogspot.co.uk/2014/07/imfw-46-vanilla-indeed.html (accesată cel mai recent la 28 martie 2019).

6. Matthew J. Salganik şi Duncan J. Watts, „Leading the herd astray: an experimental study of self-fulfilling prophecies în an artificial cultural market”, *Social Psychology Quarterly*, vol. 74, nr. 4/toamna 2008, p. 338,

DOI: https://doi.org/10.1177/019027250807100404 (accesat cel mai recent la 28 martie 2019).

7. S. Sinha şi S. Raghavendra, „Hollywood blockbusters and long-tailed distributions: an empirical study of the popularity of movies", *European Physical Journal B*, vol. 42, 2004, pp. 293 – 296, DOI: http's: //doi.org/10.1140/epjb/e2004 – 00382 – 7; http://econwpa. repec.org/eps/io/papers/0406/0406008. pdf (accesată cel mai recent la 28 martie 2019).

8. *John Carter*: analysis of a so-called flop: a look at the box office and critical reaction to Disney's early tentpole release *John Carter*", *Whatlulture*, on-line la adresa http://whatculture.com/film/john-carter-analysis-of-a-so-called-flop (accesată cel mai recent la 28 martie 2019).

9. J. Valenţi, „Motion pictures and their impact on society in the year 2000", discurs susţinut la Midwest Research Institute, Kansas City, 25 aprilie 1978, p. 7.

10. William Goldman, *Adventures in the Screen Trade*, Warner, New York, 1983.

11. Sameet Sreenivasan, „Quantitative analysis of the evolution of novelty în cinema through crowdsourced keywords", *Scientific Reports 3*, articolul nr. 2758, 2013, actualizat la 29 ianuarie 2014, DOI: http's: //doi.org/10.1038/ step02758, https://www.nature.com/articles/step02758 (accesat cel mai recent la 28 martie 2019).

12. Márton Mesty în Taha Yasseri şi János Kertész, „Early prediction of movie box office succes based on Wikipedia activity big data", *PLOS ONE*, 21 august 2013, DOI: https://doi.org/10.1371/joumal. pone.0071226 (accesat cel mai recent la 28 martie 2019).

13. Ramesh Sharda şi Dursun Delen, „Predicting box-office succes of motion pictures with neural networks", *Expert Sistems with Applications*, vol. 30, nr. 2/2006, pp. 243 – 244.

DOI: https://doi. Org/10.1016/j. eswa.2005.07.018

(accesat cel mai recent la 28 martie 2019); on-line la adresa https://www.sciencedirect.com/science/article/pii/S0957417405001399 (accesată cel mai recent la 28 martie 2019).

14. Banksy NY, „Banksy sells work for $60 în Central Park, New York – video", *Guardian*, 14 octombrie 2013, on-line la adresa https://www.theguardian.com/artanddesign/video/2013/oct/14/banksy-central-park-new-york-video (accesată cel mai recent la 28 martie 2019).

15. Bonhams, „Lot 12 Banksy: Kids on Guns", 2 iulie 2014, online la adresa http://www.bonhams.com/auctions/21829/ lot/12/ (accesată cel mai recent la 28 martie 2019).

16. Charlie Brooker, „Supposing… subversive genius Banksy is actually rubbish, *Guardian*, 22 septembrie 2006, on-line la adresa https://www.theguardian.com/commentisffee/2006/sep/22/ arts. visualarts (accesată cel mai recent la 28 martie 2019).

17. Gene Weingarten, „Pearls before breakfast: can one of the nation's greatest musicians cut through the fog of a DC rush hour? Let's fiind out", *Washington Post*, 8 aprilie 2007, online la adresa https://www.Washingtonpost.com/lifestyle/ magazine/peans-before-breakfast-can-one-of-the-nations-great-musicians-cut-through-the-fog-of-a-de-rush-hour-lets-find-out/2014/09/23/8 abd46 da-4331 – 11 e4-b47 cf5889 eoble5 fstory.html? utmterm =. a8 c9 b9922208 (accesată cel mai recent la 28 martie 2019).

18. Citatele provin dintr-o discuţie particulară purtată cu Armand Leroi. Studiul la care se referă el este: Matthias Mauch, Robert M. MacCallum, Mark Levy şi Armand M. Leroi, „The evolution of popular music: USA 1960 – 2010", *Royal Society Open Science*, 6 mai 2015, DOI: https://doi.org/10.1098/rsos.150081 (accesat cel mai recent la 28 martie 2019).

19. Citatele provin dintr-o discuţie particulară

purtată cu David Cope.

14. Acest citat a fost scurtat pentru mai multă concizie. Vezi Douglas Hofstadter, *Godel*, *Escher*, *Bach*: *An Eternal Golden Braid*, Penguin, Londra, 1979, p. 673.

15. George Johnson, „Undiscovered Bach? No, a computer wrote it", *New York Times*, 11 noiembrie 1997.

16. Benjamin Griffin şi Harriet Elinor Smith (editori), *Autobiography of Mark Twain*, vol. 3, Oakland, CA, şi Londra, 2015, partea I, p. 103.

17. Leo Tolstoy, *What Is Art*? Penguin, Londra, 1995, ediţia princeps a fost publicată în 1897.

18. Douglas Hofstadter, op. cit., p. 674.

Încheiere

1. Pentru povestea lui Rahinah Ibrahim, vezi https://www.propublica.org/article/fbi-checked-wrong-box-rahinahibrahim-terrorism-watch-list (link accesat cel mai recent la 28 martie 2019) şi http's: //alumni. stanford.edu/get/ page/magazine/article/? articleid = 66231 (link accesat cel mai recent la 28 martie 2019).

2. GenPact, *Don't underestimate importance of process* în *coming world of AI*, 14 februarie 2018, on-line la adresa http://www. genpact.com/insight/blog/dont-underestimateimportance-of-process-in-coming-world-of-ai (accesată cel mai recent la 28 martie 2019).

MULŢUMIRI

Îmi imaginez că există şi persoane cărora le este uşor să scrie. Cunoaşteţi genul – acelea care sar din pat înainte de ivirea zorilor şi care până la prânz au scris deja un capitol şi uită să coboare la cină pentru că sunt atât de absorbite de elanul lor creator, încât nu şi-au dat seama că a trecut timpul.

Eu, una, cu siguranţă nu sunt una dintre acestea.

Trecerea prin acest proces de creaţie a însemnat o luptă cotidiană cu acea parte din firea mea care nu vrea decât să stea pe canapea, ronţăind chipsuri şi uitându-se la Netflix, şi un război declarat cu valurile de panică şi de îngrijorare pe care credeam că le-am lăsat în urmă în

momentul în care mi-am terminat doctoratul. Nu pot să spun că *am scris* cartea aceasta, ci, mai degrabă, că am smuls-o din mine, lovind şi ţipând. Uneori, la propriu.

Prin urmare, sunt cu atât mai recunoscătoare remarcabilelor persoane care au fost dispuse să mă ajute pe parcurs. Extraordinarul colectiv de redacţie, care a fost atât de generos cu timpul şi ideile sale pe tot parcursul anului trecut: Susanna Wadeson, Quynh Do, Claire Conrad, Emma Parry, Gillian Somerscales, Emma Burton, Sophie Christopher, Hannah Bright, Caroline Saine şi toţi oamenii de la Janklow şi Nesbit, Transworld şi Norton care m-au ajutat din culise. Acelaşi lucru este valabil pentru Sue Rider, Kat Bee şi Tom Copson: m-aş fi rătăcit de-a dreptul fără voi!

Mulţumiri din toată inima şi celor pe care i-am intervievat, dintre care unii sunt citaţi în text, deşi toţi m-au ajutat să conturez ideile din această carte: Jonathan Rowson, Nigel Harvey, Adam Benforado, Giles Newell, Richard Berk, Sheena Urwin, Steyve Coloan, Mandeep Dhami, Adrian Weller, Toby Davies, Rob Jenkins, Jon Kanevsky, Timandra Harkness, Dan Popple şi echipa de poliţie din West Midlands, Andy Beck, Jack Stil goe, Caroline Rance, Paul Newman, Phyllis Illarmi, Armand Leoni, David Cope, Ed Finn, Kate Devlin, Shelia Hayman, Tom Chatwin, Carl Gombrich, Johnny Ryan, Jon Crowcroft şi Frank Kelly.

Îi mai adaug aici pe Sue Webb şi Debbie Enright de la Network Typing, precum şi pe Sharon Richardson, Shruthi Rao şi Will Storr, al căror ajutor a fost de nepreţuit pentru a reuşi să dăm formă cărţii de faţă. Pe lângă asta, în momentul în care am scris, în sfârşit, ceva care să semene cu nişte fraze, James Fulker, Elizabeth Adlington, Brendan Maginnis, Ian Hunter, Omar Miranda, Adam Dennett, Michael Voale, Jocelyn Bailey, Cât Black, Tracy Fry, Adam Rutherford şi Thomas Oleron Evans m-au ajutat să găsesc defectele cele mai evidente şi să le îmblânzesc. Îi mulţumesc şi lui Geoff Dahl, care, pe lângă sprijinul moral

pe care mi l-a oferit pe parcursul întregului proces, a avut şi o idee foarte bună pentru designul copertei.

Multe, multe mulţumiri comitetului meu de referenţi: Elizabeth Cleverdon, Bethany Davies, Ben Dickson, Mike Downes, Charlie şi Laura Galan, Katie Heath, Mia Kazi-Fomari, Fatah Ioualitene, Siobhan Mathers, Mabel Smaller, Ali Seyhun Saral, Jennifer Shelley, Edward Steele, Daniel Vesma, Jass Ubhi.

De asemenea, sunt nespus de recunoscătoare familiei mele pentru susţinerea neabătută şi statornica loialitate. Phil, Tracy, Natalie, Marge & Parge, Omar, Mike şi Tania – aţi avut mai multă răbdare cu mine decât am meritat adesea! (Nu o luaţi chiar *ad litteram*, totuşi, pentru că s-ar putea să mai scriu o carte şi voi avea din nou nevoie de ajutorul vostru, bine?)

Şi, în cele din urmă, dar nu în ultimul rând, multe mulţumiri lui Edith! La drept vorbind, tu nu m-ai ajutat în niciun fel, dar nici nu aş fi vrut să fie altfel.

GLOSAR

A (se) autentifica (*log on*/*în* sau *sign* în engl.) –", a deschide o sesiune de lucru cu un calculator aflat la distanţă. În mod obişnuit, pentru a realiza această operaţie sunt necesare un nume de utilizator şi o parolă" (Traian Anghel, *Dicţionar de informatică*, Editura Corint, Bucureşti, 2017). Operaţia inversă se numeşte deconectare (*log*/ *sign off*, engl.).

Adevăr întemeiat pe observaţie (*ground truth*, engl.) – informaţie empirică indubitabilă, obţinută direct, prin observare, pe care se bazează un algoritm. Sintagma este folosită în numeroase domenii (tactică militară, geografie, cartografie, meteorologie etc.).

Agent software (*software agent*, engl.) – termen folosit în informatică pentru a desemna un program de computer capabil să acţioneze (relativ) autonom şi să manifeste următoarele caracterisitei: autonomie, proactivitate, reactivitate, adaptabilitate, caracter cognitiv şi social.

Algoritm (*algorithm*, engl.) – metodă (procedură) de calcul care descrie toate instrucţiunile sau toţi paşii necesari care trebuie parcurşi într-o anumită ordine pentru a obţine un anumit rezultat.

Aplicaţie mobilă (*mobile application*/*mobile app*, engl.) – aplicaţie mobilă, program sau software menit să ruleze pe dispozitive mobile (telefon mobil, tabletă, smartwatch etc.).

Asistent (personal) virtual (*virtual assistant*/*intelligent personal assistant*, engl.) – asistent (personal) virtual, agent software, program care poate îndeplini diverse sarcini pe baza unor comenzi vocale ale utilizatorului, putând chiar să răspundă folosind o voce sintetizată. Siri a companiei Apple, Alexa a Amazonului şi Google Assistant sunt cei mai cunoscuţi asistenţi virtuali.

Bug (engl.) – „(deficienţă, hibă), eroare de proiectare sau de programare care are drept consecinţă funcţionarea incorectă a unui produs software” (Traian Anghel, op. cit.)

Cookie (*cookie HTTP*, *web cookie sau browser cookie*, *engl; temoin de connexion*, fr.) – informaţie (codificată) trimisă de un site şi stocată în computerul utilizatorului acestuia de către browserul său web, în timp ce utilizatorul accesează site-ul.

Broker de date sau de informaţii (*data broker*, engl.) – companie specializată în colectarea informaţiilor despre persoane care există în surse publice sau private (de la recensăminte la registre ale adreselor, baze de date cu înregistrările de vehicule sau cu permisele de conducere, informaţii furnizate on-line de utilizatorii diverselor site-uri, registre ale tranzacţiilor bancare, istorice ale căutărilor on-line etc.), pe care le analizează şi le însumează, pentru a realiza profiluri individuale (care pot conţine sute şi chiar mii de informaţii diverse despre o persoană sau un grup de persoane, de la gen, rasă, religie şi până la venit lunar, probleme de sănătate etc.) pe care le vând apoi organizaţiilor care doresc să organizeze campanii de reclame bine ţintite (*targeted advertising*).

Data mapping (engl.) – descoperirea corespondenţelor dintre informaţiile deţinute (*mise en correspondance de données*, fr.), termen folosit în informatică şi în managementul informaţiei pentru a defini crearea unor aplicaţii matematice menite să realizeze corespondenţe între două sisteme de date (informaţii). Reprezintă primul pas în îndeplinirea unor sarcini de integrare de date diverse.

Data mining (engl.) – exploatarea bazelor de date (*exploration de données*, fr.), extragerea de cunoştinţe din baze de date, descoperirea de patternuri, de legături noi între diversele informaţii stocate.

Date (informatizate) (*data*, engl.) – informaţii sau „elemente de informaţie" (Traian Anghel, op. cit.).

Deep leaming (engl.) – învăţare profundă (numită în limba engleză şi *deep structured leaming sau hierachical leaming*), modul în care învaţă o reţea neurală; ansamblu de metode de învăţare automată specifice reţelelor neurale, capabile să creeze modele cu un înalt nivel de abstractizare din informaţiile puse la dispoziţie, folosite cu succes în domenii precum analizarea sunetelor, recunoaştere facială şi vocală etc.

Deteriorarea relaţiilor odată cu creşterea distanţei (*distance decay*, engl.) – expresie folosită în geografie pentru a descrie modul în care distanţa afectează interacţiunile (sociale, culturale etc.) dintre două centre, care postulează că interacţiunile dintre centre scad pe măsură ce distanţa dintre ele creşte. Efectul de deteriorare a relaţiilor mai este cunoscut şi sub numele de „prima lege a geografiei", formulată de geograful şi cartograful american Waldo Tobbler, conform căreia: „toate lucrurile au relaţii cu toate celelalte, dar cele apropiate sunt mai legate unele de altele decât acelea aflate la distanţă una de alta".

Digital Obiect Identifier (DOI) – şir de litere sau semne folosit pentru a identifica un document electronic. Nu se modifică, spre deosebire de URL-uri care se

transformă de îndată ce se modifică structura serverului unde a fost publicat respectivul document on-line, aşadar transmiterea DOI este o metodă de identificare mai bună decât prezentarea adresei URL.

Fals-pozitiv – (în medicină) rezultatul unui test care atestă prezenţa unui lucru inexistent; (în statistică) eroare de tip I (care neagă ipoteza nulă, care era corectă).

Feed (engl.) – „flux de ştiri sau canal de ştiri, care furnizează informaţii de sumar despre articolele noi publicate într-un site sau blog” (Traian Anghel, op. cit.)

Filtru de internet – numit şi software de controlare a conţinutului, reprezintă un ansamblu de tehnici care vizează limitarea accesului la anumite site-uri.

Informatică (*computer Science*, engl.) – ştiinţa care studiază procesele de prelucrare automată a informaţiilor (datelor), procese care pot fi, la rândul lor, reprezentate ca informaţii (sau programe informatice) sau „modalităţile practice pentru implementarea informaţiilor în sistemele de calcul” (Traian Anghel, op. cit.). Presupune folosirea algoritmilor pentru prelucrarea, stocarea şi comunicarea informaţiilor digitale. Informatica teoretică studiază teoria limbajelor formale (automatica, teoria computaţională şi teoria complexităţii, teoria graiurilor etc.).

Ipoteza nulă – ipoteză presupusă a fi adevărată dacă nu este infirmată statistic de rezultate.

Învăţare automată (*machine leaming*, ML, engl; *apprentissage automatique*, fr.), domeniu al informaticii şi ramură a studiilor dedicate inteligenţei artificiale (IA) care cercetează algoritmii şi modelele statistice care permit sistemelor informatice să îndeplinească o sarcină în lipsa unor instrucţiuni specifice, doar pe baza decelării unor modele şi a realizării unor inferenţe logice.

Listă de corespondenţă (*mailing list*, engl.) – listă care cuprinde adresele sau e-mailurile unor clienţi potenţiali cărora, în cadrul unor strategii de marketing direct, li se trimit produse sau informaţii despre acestea (prin intermediul newsletterelor, broşurilor diverse sau

reclamelor).

Mass-marketing (engl.) – marketing de/în masă, care se adresează tuturor, fără a face nicio diferenţiere între diversele segmente de piaţă.

Maşină (*machine*, engl.) – „termen echivalent cu calculator, gazdă sau sistem” (Traian Anghel, op. cit.).

Mesagerie instantanee (MI) (desemnată şi prin anglicismul chat, pronunţat *cel*) (*instant messaging*, IM, engl.) – tehnologie care oferă transmitere instantanee de mesaj text pe internet. Skype sau Facebook Messenger sunt MI.

Micro-targetare (*micro-targeting*, engl.) – strategie folosită în general de partidele politice, pentru a trimite reclame create special pentru segmente foarte mici de alegători, care au fost grupaţi cu ajutorul unor procese de *data mining*.

Nişă de piaţă (*market niche*, engl.) – „segment de piaţă de dimensiuni reduse, sector foarte atractiv în cadrul unei pieţe, care oferă oportunităţi pentru firmă (avantaje competitive şi concurenţiale) în ceea ce priveşte vânzările, rentabilitatea, lipsa concurenţilor puternici etc”. (apud şef lucrări dr. Florin D. Frone, *Dicţionar de marketing*, *fa*.).

Prejudecată (*bias*, engl.) – predispoziţia de a face unele alegeri care nu sunt fondate pe raţionament; în cazul algoritmilor, acest termen denotă o eroare sistematică ce a fost introdusă prin intermediul unui eşantion care favorizează anumite rezultate. Spre exemplu, un algoritm de detectare a feţelor în poze ar putea avea bias către oamenii albi, întrucât eşantionul pe care a fost conceput nu conţinea exemple cu oameni de culoare.

Provider de internet (*Internet Service Provider*, ISP, engl.) – numit şi furnizor de servicii internet, reprezintă o companie care le oferă clienţilor săi conexiune şi acces la internet.

Reclame ţintite (*targeted advertising*, engl.) – formă de advertising on-line care foloseşte metode sofisticate pentru a putea să se adreseze doar persoanelor celor mai

interesate de serviciul sau produsul căruia i se face reclamă. Selectarea potenţialilor consumatori se face pe baza unor criterii diverse (gen, rasă, vârstă, statut economic, nivel de şcolarizare, comportament la shopping, istoric al cumpărăturilor on-line etc.). Reprezintă opusul publicităţii nediferenţiate, care este o strategie de *mass-marketing*.

Reţea neurală (artificială) (*artificial neural network*, *ANN*, engl.) – reţea compusă din neuroni artificiali, structură menită să proceseze informaţiile a cărei sursă de inspiraţie au constituit-o reţelele neurale biologice şi modul acestora de a structura informaţiile. Sunt capabile să înveţe din experienţă.

Secvenţierea ADN-ului (*sequencing DNA*, engl.) – procedură prin care se analizează codul ADN.

Ştiinţa datelor (*data Science*, engl.) – domeniu de studiu multidisciplinar care foloseşte diverse metode ştiinţifice pentru extragerea informaţiilor din baze de date bine structurate sau nestructurate. Concepte similare: *data mining* sau *big data*.

Thread (engl.) – fir de execuţie, „în programare, flux de instrucţiuni care se execută în interiorul unui proces […] care poate fi format din mai multe asemenea fire, care se execută în paralel, având însă în comun toate resursele principale caracteristice procesului”. (Traian Anghel, *Dicţionar de informatică*); (pe un forum) fir de discuţie, subiect comentat de mai mulţi utilizatori.

Urme digitale/informatice (*digital shadow sau digital footprint*, engl.) – totalitatea informaţiilor despre o persoană decelabile din activităţile pe care aceasta le desfăşoară în mediul on-line sau pe aplicaţiile mobile (telefonie, servicii de mesagerie etc.): acţiuni diverse, produse achiziţionate on-line, mesaje comunicate, conţinuturi apreciate etc. Deşi pot să nu prezinte un interes prea mare dacă rămân izolate, devin o sursă inestimabilă de informaţii atunci când sunt cumulate şi interpretate, căci contribuie la crearea profilului psihologic

al utilizatorului.

INDEX

23 and Me 145 – 147 52 Metro 227

Academia Naţională de Ştiinţe a SUA 221 Acxiom 52 Adele 246

AF447 (zbor) 173, 175 Afigho, Chukwuemeka 16 Al Music 246 Air France 173 – 176, 179 Alhambra 204 Alton Towers 38, 39 ALVINN 156 – 157 Alzheimer (maladie) 124 – 126, 145, 147

Amazon (motor de recomandări) 24, 229 Ancestry.com 147 Apple 73, 215 Audi 178, 183

Bainbridge, Lisanne 175 – 177, 180

Banksy (artist) 192, 237 Baril, David 222, 223

Bardett, lamie 69 Barwell, Clive 190 – 192 Bayes, teorema lui 159 – 164, 167

Bayes, Thomas 162 Beck, Andy 114, 129, 130 Bell, Joshua 238 Berk, Richard 90, 93, 99 Bonin, Pierre-Cedric 173 – 175 Bratton, Bill 194 – 195, 198 Brixton 75 – 77

Brooks, Christopher Drew 92, 93, 108

Brown, Joshua 177 Burgess, Ernest W. 82 – 83

Cambridge Analytica 62 – 67, 68, 72

CAMELYON16, 121, 122 Centaur Chess 257 China 112, 113, 217 Citroën DS19 154 Citymapper 41 Clubcard (Tesco) 45 – 46 Coloan, Steyve 202 Commodore 64 11

Compătat 195 Cope, David 242 – 244, 247 Corelogic 52 Creemers, Rogier 72 cutremure 198 – 199

Daimler 153, 171 DARPA 150 – 152, 155, 157, 158, 167

Davies, Toby 204, 205 Deep Blue 19 – 21, 23, 132, 257

Deep Mind 139, 140 Dewes, Andreas 59 Dhami, Mandeep 106, 107, 108

Dodds, Peter 226 Douglas, Neil 211, 212 Dubois, căpitanul 179 Duggan, Mark 75 Dunn, Edwina 45

ebureau 52 Eckert, Svea 59, 60 Eppink, Richard 34 – 35 Epstein, Robert 31 – 32, 68 Equivant (fosta Northpointe) 99, 272

Experimente de Inteligenţă Muzicală (EMI) 242 – 245, 247 – 248
Face ID (Apple) 215 Facebook 16, 24, 25, 43, 58, 63 – 64, 67 – 68, 73 Face First 221
Face Net (Google) 217 – 218 FBI 191, 208, 209, 212, 214, 217, 222, 252 Financial Times 153 Firebird II154 Firefox 73 Ford 66, 153, 171 Fuchs, Thomas 135 Fundaţia Clinton 66
Galton, Francis 143 – 145 GDPR (General Data Protection Regulation), 72 General Motors 154 Germania 59, 138 Goldman, William 232, 235 Google 31, 32, 36, 47, 58, 60, 100, 139, 140, 141, 170, 217, 218
Google Brain 130 Google Images 99 Google Maps 158 Google Search 23 Gogle Street-View 209 Google Translate 61 GPS 17, 29, 33, 151, 158, 159, 163, 176
Guerry, André-Michel 187 – 188
Hamm, John 134 Hammond, Philip 153 Harkness, Timandra 141 Hinton, Geoffrey 119 Hipocrate 111, 112 Hofstadter, Douglas 242, 248
Hugo, Christoph von 164, 165 Humby, Clive 45, 46, 47 Hume, David 236 Hunchlab 205, 206, 223 Huntington (boală) 147
IA (inteligenţă artificială) 26, 28, 86, 118, 119, 239, 254 IBM 19, 20, 131, 132, 134.
135.139.256
Ibrahim, Rahinah 251 – 252 Innocence Project 213 Instagram 58, 222 Internet Movie Database (IMDB) 232
Jemaah Islam 252 Jemaah Islamiyah 252 Jennings, Ken 132 Jeopardy (concurs TV) 132
133.135.256 John Carter (film) 231 Johnson, Richard 76 – 78 Jones Beach 15
Jones, Robert 29 – 30, 42 Jukebox 246
Kadoodle 31 – 32 Kahneman, Daniel 102 – 103 Kanevsky, dr. Jonathan 127, 129
Kant, Immanuel 237 Kasparov, Garry 19 – 22, 30, 42, 132, 257 Kelly, Frank 119 Kemer, Winifred 241
Kemighan, Brian 12 Killingheck 190 – 191

Larson, Steve 241 Legea lui Weber 105 – 106 Leibniz, Gottfried 236 Leroi, Armand 238 – 239 LIY – ingyun71 Lickel, Charles 132 LIDAR (Light Detection and Ranging) 157 – 159 Lista cu subiecţi strategici 206, 255
„Lockdown” (52 Metro) 227 London Bridge 222 London School of Economics (LSE) 169 Loomis, Eric 99 Lucas, Teghan 210
Maple, Jack 192 – 196 Marea Britanie 201 – 202, 205 Marx, Gary 224 McGrayne, Sharon Bertsch 161
Meehl, Paul 40 Mercedes 164 – 166 Miligarth 190 – 191 Mills, Tamara 136, 138 MIT Technology Review 135 Moses, Robert 15 Music Lab 227 Musk, Elon 180 My Heritage 147
National Geographic 147 Navlab 154 Netflix 239 – 240 New England Journal of Medicine 128
New York City 233 New York Times 50, 153, 228, 241
Newman, Paul 167 – 169, 171
OKC – upid 24 Ontario 219 Operaţiunea Lynx 187, 190 – 191
Ordinul Surorilor de Notre Dame 124 Oxbotica 167
Palantir Technologies 52 PathAI 114, 129 Petrov, Stanislav 37 – 39 Pomerleau, Dean 156 – 157 Pratt, GUI 180 Pred Pol (PRED – ictive POL – icing) 199 – 201, 203 – 206, 223
programul Crimewatch 186 Proiectul Eureka Prometheus 155
ProPublica 94
radar 157 – 159 RAND Corporation 206 război nuclear 37 Războiul Rece 37 Rhode Island 89
Robinson, Nicholas 75 – 78 Rossmo, Kim 186 – 187 Rotten Tomatoes (website) 229, 232
Royal Free NHS Trust 139 Rutherford, Adam 147
Safari (browser) 73
Sainsbury’s 46 Salganik, Matthew 226, 228 Schmidt, Eric 47 Science (revistă) 32, 166 Scunthorpe 16 Semmelweis, Ignaz 112 Sesame Credit 70, 218 Slash X (bar) 150, 152 Snowdon, David 124 Sørensen, Alan 229

Spotify 226, 229, 239 – 240 Spotify Discover 240 Sreenivasan, Sameet 232 – 233
Stammer, Neil 222 Stil goe, Jack 169 – 171 SUA 252
Talley, Steve 207 Tesco 45 – 47 Tesla 177, 181 – 182 Testul Magiei 254 Thiel, Peter 52 Timberlake, Justin 225 Timberlake, Justin (artist) 226 – 227 Tolstoi, Lev 247
Tomâom 23, 30 Toyota 180, 183 Twain, Mark 247 Twitter 25, 52, 58, 60, 64
Über 25, 178
Uniunea Libertăţilor Civile Americane (ACLU) 34 Uniunea Sovietică 224 Universitatea din Texas, Centrul de Cancer M.D. Anderson 134 URL-uri 60 – 61
Valenţi, Jack 232 Vanilla (formaţie) 229 Volvo 168, 183 Volvo XC90183
Walmart221
Walt Disney 232 Warhol, Andy 237 Washington Post 238 Waterhouse, Heidi 57 Watson (calculator IBM) 131
136, 138, 143, 256 Watts, Duncan 226 Waymo 170 – 171 Waze41 – 42 Williams, Pharrell 246 Windows XP 141 Wired 53, 177
Xing.com 60
Zaghba, Youssef 222 – 223 Zilly, Paul 91 Zuckerberg, Mark 43 ZX Spectrum 11

În curând în colecţia

CORINT FUTURE

Jamie Susskind

POLITICA VIITORULUI

Cum să trăim împreună într-o lume transformată de tehnologie

Traducere din limba engleză Adina Ihora

— Fragment – INTRODUCERE

Viitorul ne urmăreşte. Este tot timpul în aşteptare, abia ascuns vederii, pândind de după colţ sau după următorul răsărit. Nu ştim sigur ce formă va lua. Adesea ne surprinde complet nepregătiţi.

În prezent, mulţi dintre noi au impresia că se apropie o perioadă de mari frământări. Lumea pare să se schimbe

mai repede decât reuşim să o înţelegem. Ne străduim adesea să ne explicăm evenimente politice care, cu numai câţiva ani în urmă, ar fi fost de neînchipuit. Uneori, nu dispunem nici măcar de cuvintele potrivite care să le descrie. În sinea noastră ştim, însă, că acesta nu este decât începutul.

Premisa de la care porneşte cartea de faţă este că progresele neîncetate din ştiinţă şi tehnologie urmează să transforme felul în care trăim, având consecinţe profunde şi totodată înspăimântătoare în ceea ce priveşte politica. Nu suntem încă pregătiţi – intelectual, filosofic sau moral – pentru lumea pe care o creăm. În următorii câţiva zeci de ani, vechile moduri de gândire, care ne-au fost utile secole sau milenii de-a rândul, vor fi puse sub semnul întrebării. Noi dezbateri, controverse, mişcări şi ideologii vor ajunge în prim-plan. Unele dintre convingerile noastre cele mai profunde vor fi revizuite sau cu totul abandonate. Va trebui să ne imaginăm împreună, din nou, ce înseamnă să fim liberi sau egali, ce înseamnă să deţinem putere sau o proprietate, chiar şi ce înseamnă ca un sistem politic să fie democratic. Politica din viitor nu va semăna aproape deloc cu politica din trecut.

Politica secolului al XX-lea a fost dominată de o întrebare esenţială: cât de mult din viaţa noastră colectivă ar trebui să fie stabilită de stat şi cât din aceasta ar trebui lăsată în seama pieţei şi a societăţii civile? Pentru generaţia care se apropie în prezent de maturitatea politică, discuţia va fi diferită: în ce măsură ar trebui ca vieţile noastre să fie dirijate şi controlate de puternice sisteme digitale – şi în ce condiţii? Această întrebare se află în centrul cărţii *Politica viitorului*.

În următoarele câteva decenii se preconizează că vom dezvolta sisteme informatice uimitoare, dintre care unele vor rivaliza cu şi, poate, vor depăşi fiinţele umane într-o paletă largă de atribuţii, chiar şi fără să ajungă la o „inteligenţă” ca a noastră. Aceste sisteme vor fi integrate în lumea fizică, ascunse în structuri şi obiecte pe care nu

am fost obişnuiţi să le considerăm drept tehnologie. Din ce în ce mai multe informaţii despre fiinţele umane – ce facem, unde mergem, ce gândim, ce spunem, ce simţim – vor fi colectate şi înregistrate ca date, apoi vor fi sortate, stocate şi procesate digital. Pe termen lung, diferenţele dintre om şi maşină, on-line şi off-line, virtual şi real, se vor estompa.

Această transformare va aduce civilizaţiei câteva mari beneficii. Vieţile noastre vor fi îmbogăţite de noi modalităţi de a ne juca, de a lucra, a călători, a face cumpărături, a învăţa, a crea, a ne exprima, a păstra legătura cu ceilalţi, a întâlni persoane necunoscute, a coordona acţiuni, a ne menţine în formă şi a găsi înţelesuri. Pe termen lung, s-ar putea să reuşim să ne perfecţionăm mintea şi trupul, făcându-le de nerecunoscut şi eliberându-ne de limitările biologiei umane.

În acelaşi timp, însă, unele tehnologii vor ajunge să deţină o mare putere asupra noastră. Vor putea să ne forţeze să ne comportăm într-un anumit fel, cum ar fi (ca să luăm un exemplu simplu) vehiculele autonome, care refuză efectiv să depăşească limita de viteză prestabilită. Alte tehnologii vor fi puternice datorită informaţiilor pe care le colectează despre noi. Simplul fapt că ne ştim supravegheaţi ne face mai puţin predispuşi să înfăptuim lucruri percepute ca ruşinoase, păcătoase sau greşite. Alte tehnologii vor filtra lumea pe care o vedem, stabilind ce să cunoaştem, modelând felul în care gândim, influenţând ceea ce simţim şi determinând astfel modul în care acţionăm.

Cei care controlează aceste tehnologii incredibile ne vor controla din ce în ce mai mult şi pe noi. Ei vor deţine *putere*, ceea ce înseamnă că vor dispune de capacitatea constantă şi extinsă de a ne determina să facem lucruri importante, pe care, altfel, nu le-am face. Din ce în ce mai mult, ei ne vor limita *libertatea*, stabilind ce este permis şi ce este interzis. Vor stabili viitorul *democraţiei*, ducând la dezvoltarea sau la degradarea acesteia. Iar algoritmii lor

vor lua decizii în probleme vitale de *justiţie socială*, alocând beneficii şi împărţindu-ne pe ierarhii, în funcţie de statutul nostru şi de respectul pe care îl inspirăm.

În consecinţă, autorităţile politice – statele, în general – vor dispune de mai multe instrumente de control decât oricând, iar marile companii tehnologice se vor bucura, la rândul lor, de putere la un nivel care va eclipsa orice altă entitate economică din timpurile moderne. Pentru a face faţă acestor noi provocări, vom avea nevoie de modernizarea radicală a ideilor noastre politice. […]

URMĂTOAREA MARE SCHIMBARE

Trăim deja într-o perioadă de profundă nelinişte politică, în fiecare zi aflăm ştiri despre războaie civile sângeroase, deplasări masive ale populaţiei, naţionalism etnic, violenţă sectară, extremism religios, schimbări climatice, tulburări economice, globalizare derutantă, inegalitate crescândă şi o diversitate de alte evenimente problematice, prea urâte ca să le amintim aici. S-ar părea că lumea nu o duce prea bine şi că discursurile noastre publice nu sunt la înălţimea aşteptărilor. Se manifestă o amplă neîncredere şi un mare dispreţ faţă de elitele politice. Două recente exerciţii democratice de masă din lumea vorbitoare de limbă engleză, alegerile prezidenţiale din SUA din 2016 şi referendumul privind Brexit din Marea Britanie, au fost pline de resentimente, chiar şi după standardele de nefericire obişnuite, facţiunile opuse străduindu-se nu numai să-şi învingă rivalii, ci şi să-i nimicească. Ambele au fost câştigate de partea care promitea să distrugă vechea ordine. Niciunul dintre aceste exerciţii nu a adus după sine împăcare sau mulţumire. Din ce în ce mai mult, după cum remarca Barack Obama la sfârşitul mandatului său prezidenţial, „totul este adevărat şi nimic nu este adevărat”. Devine din ce în ce mai dificil pentru cetăţenii obişnuiţi (indiferent de apartenenţa lor politică) să facă distincţia dintre adevăr şi impostură, realitate şi zvon, semnal şi zgomot. Mulţi chiar au renunţat să mai încerce. Suntem tentaţi să rămânem pe poziţii şi să

ţinem piept furtunii din prezent, fără să ne gândim prea mult la viitor.

Însă aceasta ar fi o greşeală.

Dacă cele mai importante previziuni despre viitorul tehnologiei se apropie de adevăr, atunci schimbarea care se întrevede ar putea fi cel puţin tot atât de importantă pentru umanitate precum revoluţia industrială, revoluţia agrară sau chiar inventarea limbii vorbite. Prin comparaţie, multe dintre problemele actuale vor fi eclipsate. Gândiţi-vă la efectul pe care l-a avut deja tehnologia în viaţa noastră – asupra felului în care lucrăm, comunicăm, ne tratăm afecţiunile, facem exerciţii fizice, mâncăm, învăţăm şi socializăm –, iar apoi amintiţi-vă că, din perspectivă istorică, era digitală are o vârstă de numai câteva secunde. În total, 99, 5% din existenţa umană s-a petrecut în epoca paleolitică, apărută cu aproximativ trei milioane de ani în urmă, când oamenii au început să folosească unelte primitive. Epoca aceea a luat sfârşit în urmă cu vreo 12.000 de ani, odată cu ultima glaciaţiune. În timpul acestei lungi perioade crepusculare, oamenii nu au remarcat aproape niciun fel de schimbare culturală. „Lumea în care oamenii intrau la naştere era aceeaşi cu cea pe care o părăseau în momentul morţii”. Dacă vă gândiţi că cele mai vechi civilizaţii umane au apărut în urmă cu vreo 5.000 de ani, atunci cei şaptezeci şi ceva de ani pe care i-am trăit alături de calculatoarele moderne, cei treizeci şi ceva de ani de când dispunem de World Wide Web şi cei zece ani de când utilizăm telefoanele inteligente nu par nicidecum nişte perioade îndelungate. Pe când timpul se scurge în mod liniar, numeroasele progrese din tehnologia digitală au loc exponenţial, viteza schimbării fiind, de la un an la altul, tot mai mare.

Nu avem niciun indiciu cu privire la viitor, aşadar încercarea de a-l anticipa este inerent riscantă şi dificilă. Îi admir pe cei care se străduiesc să facă acest lucru în mod riguros şi m-am inspirat pe larg din activitatea lor la întocmirea acestei cărţi. Dar ca să fim realişti, ar trebui să

începem prin a recunoaşte că astfel de predicţii se înşală adesea foarte tare. Probabil că o mare parte a viitorului anticipat în paginile următoare nu va deveni niciodată realitate şi alte evenimente, complet neprevăzute, îşi vor face în schimb apariţia, luându-ne prin surprindere. Acestea fiind spuse, cred că se pot face supoziţii de bun-simţ, în cunoştinţă de cauză, despre felul în care ar putea arăta viitorul în baza a ceea ce cunoaştem astăzi despre tendinţele actuale din ştiinţă, tehnologie şi politică. Riscul cel mai mare ar fi să nu încercăm în niciun fel să anticipăm viitorul.

Se povesteşte despre o întâlnire dintre omul de stat din epoca victoriană William Gladstone şi sclipitorul om de ştiinţă Michael Faraday. Acesta din urmă încerca să-i explice lui Gladstone activitatea lui revoluţionară în domeniul electricităţii, dar Gladstone nu părea impresionat.

— Dar la ce este bună chestia asta? a întrebat cel din urmă, cu o frustrare crescândă. La *ce foloseşte* ea?

— Ei bine, domnule, i-a răspuns Faraday, ajuns la limita răbdării, sunt şanse mari ca, în curând, să puteţi pune un impozit pe asta.

Multor inovatori, asemenea lui Faraday, le este greu să explice implicaţiile sociale şi practice ale activităţii lor. Iar noi, ceilalţi, asemenea lui Gladstone, respingem mult prea adesea tehnologiile pe care încă nu le înţelegem. Poate fi greu de observat importanţa politică a invenţiilor care, la prima vedere, par să nu aibă nimic în comun cu politica. Când ne aflăm în faţa unui *gadget* sau a unei aplicaţii noi, nu ne gândim mai întâi la implicaţiile politice pe care le vor avea acestea. Vrem, în schimb, să ştim: la ce foloseşte asta? Cât costă? De unde îmi pot procura şi eu aşa ceva? Nu este deloc surprinzător. În general, ne întâlnim cu tehnologia în calitate de consumatori, însă această atitudine relativ obtuză trebuie să se schimbe. Trebuie să dedicăm noilor tehnologii ale puterii aceeaşi examinare atentă şi acelaşi scepticism pe care le-am

rezervat întotdeauna politicienilor puternici. Tehnologia nu ne afectează numai în calitate de consumatori, ci şi de cetăţeni. În secolul al XXI-lea, *digitalul este politizat*.

Această carte li se adresează, pe de o parte, acelor Gladstone care vor să înţeleagă mai mult din tehnologie, iar pe de altă parte, acelor Faraday care vor să înţeleagă mai bine importanţa politică a activităţii lor. Însă cea mai mare parte a cărţii se adresează oamenilor obişnuiţi, care doresc să priceapă viitorul ceva mai bine, dacă nu din alte motive, măcar pentru a-i responsabiliza pe toţi acei Gladstoni şi Faraday.

INGINERII FILOSOFI

Reflectaţi puţin la fragmentul de mai jos:

Cinste celor nebuni. Celor inadaptaţi. Celor rebeli. Celor zurbagii. Cuielor rotunde din găurile pătrate. Celor care văd lucrurile altfel. Celor cărora nu le plac regulile. Şi care nu au niciun respect faţă de starea actuală a lucrurilor. Poţi să-i citezi, să nu fii de acord cu ei, să-i lauzi sau să-i denigrezi. Singurul lucru pe care nu-l poţi face este să-i ignori. Pentru că ei schimbă lucrurile. Fac rasa umană să avanseze. Deşi unii i-ar putea considera nebuni, noi vedem în ei nişte genii. Pentru că oamenii care sunt destul de nebuni să creadă că pot să schimbe lumea sunt cei care chiar o schimbă.

Acestea nu sunt cuvintele unui politician. Sunt extrase din postsincronul de la „Think Different”[89], o reclamă a companiei Apple din 1997, în care erau incluse imagini ale unor rebeli celebri, printre care Mahatma Gandhi şi Martin Luther King. Reclama prezintă o anumită viziune asupra lumii, larg răspândită printre antreprenorii din domeniul tehnologiei, şi anume că activitatea lor are şi o importanţă filosofică, pe lângă cea comercială. „În Silicon Valley este ceva obişnuit”, arată Jaron Lanier[90], „ca persoanele foarte tinere, cu un start-up amenajat în vreun

89 Gândeşte diferit (n. tr.)

90 Jaron Lanier (n. 3 mai 1960, New York, SUA), programator, muzician şi părintele realităţii virtuale (n. tr.).

garaj, să declare că ţelul lor este să schimbe din temelii cultura umană, la scală mondială, în numai câţiva ani, şi că nu sunt încă pregătiţi să-şi facă griji în privinţa banilor pentru că acumularea unei mari averi este o chestiune măruntă, care se va rezolva de la sine”. Un astfel de mod de gândire are ceva atrăgător, în parte pentru că sugerează că aceste companii tehnologice ar putea să nu fie atât de rapace pe cât sunt considerate îndeobşte. Iar premisa de la care se porneşte este una corectă: tehnologiile digitale chiar *au o* capacitate uimitoare de a schimba lumea. Comparaţi afirmaţiile următoare:

„Filosofii nu au făcut decât să *interpreteze* lumea în diferite moduri; ideea este, însă, *schimbarea* sa”.

„Noi nu analizăm o lume, e\ o făurim”.

Prima datează din 1845 şi îi aparţine lui Karl Marx – *Teze despre Feuerbach*. A servit drept slogan revoluţionarilor politici mai mult de o sută de ani de la publicarea sa. Cea de-a doua îi aparţine manieratului Tim Berners-Lee, inventatorul

World Wide Web. Marx şi Berners-Lee nici că ar putea fi mai diferiţi în ceea ce priveşte gândirea politică, temperamentul sau stilul lor de păr facial. Însă ceea ce au în comun – în plus faţă de faptul că au schimbat mersul istoriei omenirii – este convingerea că există o deosebire între *a aduce o* schimbare şi doar a te *gândi* la schimbare sau a *studia* schimbarea. Din această perspectivă, departe de a fi o prezenţă spectrală asupra căreia nu avem niciun control, viitorul este ceva pe care îl elaborăm şi îl construim.

„Noi nu suntem filosofi experimentalişti”, spunea cândva Berners-Lee, „suntem ingineri filosofi”. Este un mod practic şi participativ de a privi viaţa; unul mai familiar constructorilor şi inventatorilor, decât cadrelor universitare îmbrăcate în stofa tweed sau filosofilor cu pulovere pe gât. Se întâmplă totodată să fie şi mentalitatea definitorie a epocii în care trăim. În prezent, cele mai importante revoluţii nu au loc în departamentele de

filosofie, nici chiar în parlamente sau în agora oraşelor, ci în laboratoare, centre de cercetare, companii tehnologice şi centre de date. Majoritatea presupun dezvoltări ale tehnologiilor digitale. Şi, totuşi, aceste avansuri extraordinare se petrec într-un alarmant climat de izolare culturală şi intelectuală. Cu puţine excepţii, între arte şi ştiinţă se cască o prăpastie. Filosofia politică şi politica socială apar rareori în programele de studii universitare din domeniul ştiinţei, tehnologiei, ingineriei şi matematicii. Iar dacă îl întrebi pe studentul clasic la arte liberale cum funcţionează un calculator, este puţin probabil să obţii un răspuns elaborat.

În companiile tehnologice, extrem de puţini ingineri sunt însărcinaţi să se gândească intens la consecinţele sistemice ale activităţii lor. Celor mai mulţi li se dau spre rezolvare probleme tehnice concrete. În sectorul tehnologic, inovaţia este, până la urmă, motivată de obţinerea de profit, chiar dacă investitorii sunt pregătiţi să adopte o abordare de genul „o idee bună la început, iar profitul mai târziu”. Nu este o observaţie critică: înseamnă doar că nu există niciun motiv pentru care a face bani şi a face lumea mai bună să însemne întotdeauna unul şi acelaşi lucru. De fapt, aşa cum o arată multe dintre exemplele din această carte, există o sumedenie de dovezi care sugerează că tehnologia digitală este mult prea adesea concepută din perspectiva celor puternici şi privilegiaţi.

Odată cu trecerea timpului, vom avea nevoie de mai mulţi ingineri filosofi demni de acest nume. Şi va deveni încă şi mai important pentru noi, ceilalţi, să interpretăm în mod critic activitatea companiilor de tehnologie, nu în ultimul rând deoarece cultura organizaţională din domeniul tehnologiei este recunoscută pentru lipsa ei de diversitate. Aproximativ nouă din zece directori executivi din Silicon Valley sunt bărbaţi. În ciuda faptului că afro-americanii alcătuiesc cam 10% din absolvenţii din IT şi 14% din forţa de muncă totală, aceştia ocupă sub 3% din

posturile din domeniul informaticii în Silicon Valley. Şi numeroase persoane din comunitatea tehnologică nutresc vederi politice ferme, care se situează mult în afara curentului general. Peste 44% din cei care au adoptat moneda Bitcoin în anul 2013, de exemplu, s-au declarat a fi „libertarieni sau anarho-capitalişti care se pronunţă în favoarea eliminării statului”.

Aşa cum voi căuta să arăt în continuare, riscăm foarte mult atunci când delegăm chestiuni de importanţă politică grupului infim de persoane care se întâmplă să fie însărcinate cu dezvoltarea tehnologiilor digitale la un moment dat. Acest lucru este adevărat indiferent dacă îi admiri pe inginerii filosofi din Silicon Valley sau dacă eşti de părere că cei mai mulţi „fraţi întru tehnologie” au rafinamentul politic al unui tranzistor. Avem nevoie de un cadru intelectual care să ne ajute să ne gândim limpede şi cu simţ critic la consecinţele politice ale inovării digitale. Cartea aceasta speră să contribuie la un astfel de cadru, utilizând ideile şi metodele teoriilor politice.

PROMISIUNEA TEORIILOR POLITICE

Scopul filosofiei, considera Isaiah Berlin, este întotdeauna acelaşi: să-i ajute pe oameni „să se înţeleagă pe ei înşişi şi astfel să acţioneze deschis, nu la întâmplare, nu în necunoştinţă de cauză”. Acesta este şi scopul nostru. Teoria politică îşi propune să înţeleagă politica prin intermediul conceptelor pe care le utilizăm pentru a ne referi la aceasta. Ce este *puterea*? Când ar trebui restricţionată *libertatea* şi pe ce considerente? *Democraţia* cere ca toată lumea să posede *aceeaşi* abilitate de a modela procesul politic? Ce înseamnă o distribuire *corectă* a resurselor societăţii? Iată genul de întrebări la care teoreticienii politici încearcă să găsească un răspuns. Această disciplină are o istorie deopotrivă îndelungată şi bogată. De la Platon şi Aristotel din Academia din Grecia antică, până la Thomas Hobbes şi Jean-Jacques Rousseau din tumultul Europei moderne timpurii, dar şi titanii gândirii politice ai secolului al XX-lea precum Hannah

Arendt şi John Rawls, gânditorii politici occidentali au încercat să clarifice şi să observe critic lumea din jurul lor, întrebându-se de ce era în felul acela şi dacă nu cumva ar putea sau ar trebui să fie diferită.

Din câteva considerente, teoria politică este cât se poate de adecvată pentru analizarea interacţiunii dintre tehnologie şi politică. În primul rând, canoanele gândirii politice conţin înţelepciunea care a supravieţuit civilizaţiilor. Ea poate pune în lumină situaţii dificile viitoare şi ne poate ajuta să ne dăm seama ce este în joc. Ar fi o nesăbuire din partea noastră să nu profităm de tezaurul ideilor pe care le avem la dispoziţie, chiar dacă, în cele din urmă, vom decide că unele din ele trebuie actualizate sau „resetate”. Teoriile politice ne oferă totodată *metode* de gândire despre lume, care ne ajută să ridicăm nivelul dezbaterilor deasupra aserţiunilor şi a prejudecăţilor.

În opinia mea, lucrul cel mai bun în privinţa teoriilor politice este că acestea abordează temele şi întrebările *majore* ale politicii. Ele oferă o vedere panoramică a pădurii politice, acolo unde alte abordări s-ar putea pierde printre copaci (sau s-ar putea înţepeni printre ramuri). În cazul nostru, acest lucru este necesar pentru a respecta tema cărţii de faţă. În situaţia în care considerăm că tehnologia ar putea avea un impact fundamental asupra condiţiei umane, atunci şi analiza acestui impact ar trebui să fie una fundamentală. Iată de ce, în cartea de faţă, ne referim la patru concepte politice fundamentale: Puterea: Cum îi domină cei puternici pe cei slabi Libertatea: Ce este permis şi ce este interzis Democraţia: Cum pot oamenii să guverneze Justiţia socială: Ce îndatoriri avem unii faţă de ceilalţi în perioade de mari schimbări, sugerez eu, merită să ne întoarcem la principiile iniţiale şi să ne gândim la aceste concepte în afara unui anumit regim juridic. Am reuşi, astfel, să ne imaginăm un sistem superior celui pe care l-am moştenit.

Teoriile politice sunt utile şi pentru că ne permit să

reflectăm critic nu numai cu privire la politică, ci şi despre cum *gândim* şi *vorbim* despre aceasta. Conceptele sunt „gaura cheii prin care vedem şi percepem inevitabil realitatea”. Dacă doresc să-i vorbesc despre politică vecinei mele, nu sunt nevoit să încep de la zero. Ştiu că dacă afirm despre un anumit proces că este „nedemocratic”, ea va şti destul de bine ce am vrut să spun şi ce conotaţii am dorit să transmit, fără să fiu nevoit să-i explic ce este democraţia şi de ce ar trebui ca aceasta să fie considerată un lucru bun. Asta, pentru că facem parte din aceeaşi comunitate lingvistică şi împărtăşim o „rezervă comună de noţiuni” desprinse din istoria şi mitologia noastră comună. Este un lucru convenabil.

Pe de altă parte, ce dorim să exprimăm despre politică poate fi uneori limitat de puţinătatea cuvintelor de care dispunem. Unele lucruri par inexprimabile sau de neconceput pentru că rezerva comună de noţiuni încă nu s-a dezvoltat suficient pentru a le da glas. „Limitele limbajului meu”, spunea Ludwig Wittgenstein, „sunt şi limitele lumii mele”.

În termeni politici, aceasta înseamnă că, chiar dacă am putea să vedem limpede viitorul, s-ar putea să nu dispunem de cuvintele adecvate pentru a-l descrie. Este motivul pentru care atât de des ne limităm viziunea asupra viitorului, la o versiune supraaccelerată a lumii în care trăim deja. „Dacă aş fi întrebat oamenii ce îşi doresc”, spunea Henry Ford, primul producător de automobile în serie, „ei mi-ar fi răspuns că vor cai mai rapizi”. Ford a recunoscut că poate fi greu să concepi un sistem radical diferit faţă de cel existent. Incapacitatea de a ne actualiza limbajul face acest lucru încă şi mai dificil.

DE CE AM SCRIS ACEASTĂ CARTE

Prima oară când am dat cu adevărat de gustul teoriei politice a fost în facultate, unde m-am îndrăgostit de această disciplină sub îndrumarea atentă a unor profesori minunaţi. S-a trezit în mine o obsesie care nu m-a părăsit nici acum. (Recunosc, retrospectiv, că cea mai reuşită

poveste de iubire din timpul studenţiei mele se poate să fi fost o idilă pasională, dar condamnată la eşec, cu filosoful german G.W.F. Hegel.)

Lăsând pasiunea la o parte, ceva din disciplina „teoriilor politice” mă răscolea. Teoreticienii politici păreau să fie mândri de gândirea lor profundă asupra *istoriei* ideilor politice, însă, cu câteva excepţii, erau aproape complet neinteresaţi de *viitorul* acestora. Mi s-a părut ciudat: de ce aceiaşi erudiţi – atât de sensibili la context când scriau despre trecut – discutau despre politică de parcă lumea din anul 2050 avea să fie la fel ca aceea din 1950? Se părea că o mare parte a unor teorii politice foarte inteligente avea o aplicare practică redusă, pentru că nu se potrivea cu realităţile din prezent. Când am reflectat la politica din viitor, gândul m-a dus la Orwell, Huxley, Wells – toţi, romancieri de la începutul secolului al XX-lea – mai degrabă decât la teoreticienii politici din secolul al XXI-lea. Se pare că nu eram singurul: de la alegerea lui Donald Trump la preşedinţia SUA, spre sfârşitul anului 2016, romanul lui Orwell, *1984*, a revenit pe lista cărţilor bestseller. Însă apare întrebarea: dacă vrem să înţelegem lumea aşa cum va fi ea în anul 2050, chiar ar trebui să ne bazăm pe o lucrare de ficţiune scrisă în 1949?

După ce am absolvit facultatea şi m-am implicat în propriile cauze politice modeste, supărătoarea mea stare de nelinişte cu privire la faptul că teoria politică ar putea fi incapabilă sau refractară să abordeze provocările care se prefigurează ale generaţiei mele a devenit o preocupare mai urgentă. Dacă tehnologia ar evolua atât de repede, încât ne-ar lipsi dotarea intelectuală să o mai înţelegem? Dacă, pe negândite, am fi pe punctul de a dezlănţui un viitor pe care nu l-am putea înţelege şi, cu atât mai puţin, controla?

Doream nişte răspunsuri, motiv pentru care am început să lucrez la cartea aceasta.

POLITICA TEHNOLOGIEI

Înainte de a trece mai departe, să începem cu o întrebare simplă: care este legătura dintre tehnologia digitală şi politică?

Tehnologia în general

Noile tehnologii *permit* realizarea unor lucruri care mai înainte nu puteau fi înfăptuite, iar acestea *fac mai uşoară* înfăptuirea unor lucruri pe care le realizam deja. Aceasta este semnificaţia lor socială de bază. Cel mai adesea, noile oportunităţi create de tehnologie sunt de natură minoră: un mod ingenios de măcinare a boabelor de cafea, de exemplu, este puţin probabil să ducă la răsturnarea unui stat. Uneori, însă, consecinţele pot fi importante. În cazul revoluţiei industriale, inventarea războaielor de ţesut mecanice şi a maşinilor de tricotat şi de filat ameninţa să elimine locurile de muncă ale pricepuţilor muncitori textilişti. Unii dintre aceştia, cunoscuţi drept „ludiţi”, au pornit o revoltă violentă în regiunea rurală a Angliei, distrugând toate noile maşini pe care le întâlneau în drumul lor. Încă mai folosim termenul de „ludiţi” pentru a-i descrie pe cei care opun rezistenţă la apariţia unor tehnologii noi şi perturbatoare.

Consecinţele *economice* ale inovaţiei, ca în cazul ludiţilor, vor impune adesea şi o reacţie politică. Însă noile tehnologii pot să aducă şi provocări de natură *morală*. Peste câţiva ani, de exemplu, sistemele de realitate virtuală vor facilita comportamente care mai înainte erau de domeniul ştiinţifico-fantasticului, inclusiv (bunăoară) posibilitatea de a face sex virtual prin intermediul unor platforme capabile de stimulare a activităţii sexuale umane. Acest lucru ridică nişte întrebări interesante. Ar trebui ca legea să permită să faci sex virtual cu un partener digital astfel conceput încât să arate ca o persoană din viaţa reală – o celebritate, de exemplu – fără cunoştinţa sau consimţământul persoanei respective? S-ar putea ca oamenii să manifeste sentimente puternice în această privinţă. Ce ar fi dacă o femeie ar face „sex” virtual cu încarnarea virtuală a prietenului celui mai bun al

soţului ei? Ar părea greşit ca legea să permită aşa ceva, însă ai putea susţine că ceea ce faci în intimitatea lumii tale virtuale, fără să produci niciun rău nimănui, nu este treaba altcuiva şi cu atât mai puţin a legii. Ca să dăm un exemplu şi mai puternic: ce-ar fi să faci sex virtual cu avatarul unui copil, în condiţiile în care nu este molestat niciun minor în cursul acelei experienţe?

Acestea sunt nişte întrebări noi. Nu a mai trebuit să răspundem până acum la ele. Poate că v-aţi format deja nişte opinii în privinţa lor. Poate că acele opinii diferă de ale mele. Exact asta şi este ideea: sunt nişte întrebări inerent *politice* şi ar trebui să li se răspundă cu atenţie, cu referire la o serie acceptabilă de principii morale. Aşa cum a remarcat un distins autor, noile tehnologii ne pot determina să ne reexaminăm vederile politice, tot aşa cum un preparat nou din meniul restaurantului preferat ne poate determina să ne reevaluăm preferinţele culinare.

Unele tehnologii sunt „inerent politice” prin aceea că necesită, de fapt, „anumite tipuri de relaţii politice” sau măcar sunt compatibile cu acestea. Când Langdon Winner scria, prin anul 1980, cita drept exemplu puterea nucleară: dacă urmează să adoptaţi un sistem bazat pe energia nucleară, atunci trebuie să acceptaţi şi faptul că va trebui ca o „elită tehnico-ştiinţifică-industrial-militară” puternică să-i administreze şi să-i supravegheze instalaţiile şi infrastructura.

Mai există şi alte tehnologii, care nu sunt inerent politice, dar care sunt politizate de contextul în care funcţionează. În Marea Britanie, de exemplu, o armă de foc autorizată este, de regulă, o tehnologie utilizată la vânarea animalelor sălbatice. Armele nu fac parte din cultura dominantă, iar oamenii sunt, în general, foarte fericiţi cu ideea unei reglementări stricte a acestora. Numai că, în SUA, dreptul de a purta arme este garantat de cel de-al doilea Amendament [al Constituţiei – n.tr.]. Opoziţia culturală faţă de această reglementare este mult mai tranşantă. Aceeaşi tehnologie, un context politic

diferit.

O ultimă legătură, mai subtilă, între tehnologie şi politică este aceea că invenţiile au modul lor de a se insinua în viaţa politică şi intelectuală. Un bun exemplu în acest sens este ceasul mecanic. Aşa cum arată Otto Mayr, unele civilizaţii antice şi-au imaginat statul ca pe un corp uman, fiecare membru component al comunităţii politice alcătuind extremităţi precum o „mână" sau un „picior". În Renaşterea târzie, metaforei corpului uman i s-au alăturat şi altele, cum ar fi „corabia statului". După Copernic, monarhul a ajuns să fie văzut ca un mare soare în jurul căruia gravitau servitorii săi. Între secolele al XVI-lea şi al XVIII-lea, metafora dominantă a fost aceea a ceasului, o maşinărie ingenioasă care inspira o „veneraţie fără precedent". După un timp, gânditorii au ajuns să vadă politica din perspectiva mecanismului unui ceas. Armonia, cumpătarea şi consecvenţa au devenit idealurile politice predominante. Oamenii tânjeau după un „om de stat-inginer atotvigilent" care să poată să „prevadă, dar şi să remedieze toate problemele posibile". Aşa se face că o anumită tehnologie şi o serie de valori politice au mers mână în mână, atingând apoteoza în secolul al XVII-lea, odată cu „extraordinara sa producţie de ceasuri" şi cu „înflorirea vădită a concepţiei autoritare despre ordine".

Şi imaginaţia politică a secolului al XX-lea a fost nuanţată de tehnologie. Chiar şi anterior erei digitale, perspectiva unor calculatoare atotputernice a inspirat o mare parte a artei şi a literaturii de ficţiune. În *Maşina se opreşte* (1928), E.M. Forster a zugrăvit o lume în care oamenii erau subordonaţi Maşinii, un sistem tehnologic global care supraveghea şi controla fiecare aspect al existenţei umane.

Unii scriitori au încercat să definească ceea ce considerau a fi ideologia timpurilor noastre. Evgheni Morozov, de exemplu, a scris despre „Doctrina Google" („încrederea entuziastă în forţa eliberatoare a tehnologiei, însoţită de impulsul irezistibil de a înrola start-up-urile din

Silicon Valley în lupta mondială pentru libertate”); „cyber-utopismul” („convingerea naivă în natura emancipatoare a comunicării on-line, care se bazează pe refuzul încăpăţânat de a-i recunoaşte dezavantajele”) şi „soluţionismul” („reformularea tuturor situaţiilor sociale complexe fie ca probleme clar definite cu soluţii clare, estimabile, fie ca procese transparente şi de la sine înţelese care pot fi optimizate cu uşurinţă – cu condiţia aplicării algoritmilor potriviţi”).

Opinia mea personală este că tehnologiile în chestiune sunt de dată prea recentă pentru a şti ce efect de durată vor avea asupra gândirii noastre politice. A fost nevoie de sute de ani pentru ca ideea ceasului mecanic să se impregneze în viaţa politică şi intelectuală a Europei. Cum alegem să interpretăm tehnologiile vremurilor noastre şi cum ne modelează acestea, la rândul lor, percepţia asupra lumii sunt chestiuni care urmează încă a fi determinate.

Tehnologiile digitale

Cam atât, pentru moment, despre relaţia generală dintre tehnologie şi politică. Această carte se referă, însă, la tehnologiile *digitale* – care mai înainte erau numite tehnologii ale informaţiei şi comunicaţiilor sau TIC. Şi se pare că acestea nu sunt numai inerent politice; sunt hiperpolitice. Asta, pentru că vizează două componente esenţiale ale vieţii politice: *comunicaţiile* şi *informaţia*.

Întreaga ordine politică se întemeiază pe coordonare, cooperare sau control. Este imposibil să organizezi viaţa colectivă fără cel puţin una din acestea trei. Şi niciuna dintre ele nu este posibilă în lipsa unui sistem de schimb de informaţii fie între oamenii obişnuiţi, fie între cei care conduc şi cei care sunt conduşi. De aceea este atât de important *limbajul*. După cum spune James Farr, fără limbaj, politica nu ar fi numai „imposibil de descris”, ci şi de-a dreptul „imposibilă”:

Statele emergente nu şi-ar putea declara independenţa, liderii nu şi-ar putea instrui adepţii,

cetăţenii nu ar putea protesta împotriva războaielor, iar instanţele nu ar putea să condamne infractorii. Nici noi nu am putea să criticăm, pledăm, promitem, susţinem, îndemnăm, cerem, negociem, să cădem la înţelegere, să facem compromisuri, să consiliem, informăm, chestionăm, sfătuim sau să consimţim. Să ne imaginăm politica fără aceste acţiuni ar însemna să ne imaginăm o politică de nerecunoscut.

Yuval Noah Harari observă, în mod elegant, că limbajul a jucat un rol esenţial în perioada timpurie a politicii umane. Un număr redus de sunete şi semne le-a permis strămoşilor noştri să emită un număr infinit de propoziţii, fiecare având o anumită semnificaţie. Aceasta le-a oferit posibilitatea să vorbească unii despre ceilalţi, să vorbească despre lucruri importante şi complicate din lumea din jurul lor, dar şi despre lucruri lipsite de orice consistenţă materială, precum miturile şi poveştile. Astfel de legende ţin comunităţile unite şi în zilele noastre.

Este posibil ca scrierea să fi apărut ca un mod de reflectare a realităţii, „însă a devenit, treptat, un mod puternic de reconfigurare a realităţii”. *Domesday Book*[91], elaborată în anul 1086, a fost realizată la cererea lui William Cuceritorul pentru a se afla întinderea şi valoarea proprietăţilor oamenilor, în vederea impozitării mai eficiente a acestora. Conţinutul registrului era „bătut în cuie” şi nu putea fi contestat. În 1179, un comentator scria despre acest „sever şi cumplit ultim bilanţ”, că:

decizia sa nu poate fi anulată sau respinsă fără teama de consecinţe. De aceea am şi numit această carte „Judecata de Apoi” (…), pentru că hotărârile sale, ca şi cele de la Judecata de Apoi, sunt irevocabile.

Nu conta decât ceea ce era trecut în registru. Puţine lucruri s-au schimbat de atunci. După cum spune Harari: „Oricine a avut vreodată de-a face cu autorităţile fiscale,

91 Domesday Book („Cartea Judecăţii de Apoi”) este un registru al anchetei privind terenurile şi animalele din cea mai mare parte a Angliei şi o parte din Ţara Galilor. în prezent se află la Arhivele Naţionale, Kew, Richmond upon Thames, în sud-vestul Londrei, (n. tr.)

sistemul de educaţie sau cu orice alt sistem birocratic stufos ştie că adevărul abia dacă mai contează. Ce apare scris în formularul tău este cu mult mai important”.

Limbajul, însă, nu este suficient. Şi comunităţile politice avansate trebuie să fie capabile să proceseze un volum mare de *informaţii* – de la rata sărăciei şi a creşterii economice, până la nivelurile fertilităţii, şomajului şi imigraţiei. Dacă vrei să guvernezi o naţiune, trebuie mai întâi să o cunoşti.

Revoluţionarii secolului al XVIII-lea, care au încercat să înlocuiască *ancien régimes*[92] decadente, au înţeles bine acest lucru. Revoluţia Franceză din 1789 a fost urmată de un efort intens de standardizare şi raţionalizare, de introducere a unui sistem unic de greutăţi şi măsuri, de împărţire a teritoriului în *departaments*[93] şi de promulgare a Codului Civil. Peste Atlantic, Părinţii Fondatori ai Statelor Unite au prevăzut necesitatea unui recensământ decenal (o „Enumerare”) chiar în primul articol al Constituţiei. Alexander Hamilton considera că guvernul federal ar trebui să constituie „centrul de informare”, pentru că astfel ar fi cel mai în măsură să „înţeleagă amploarea şi urgenţa pericolelor care ameninţă să-şi facă apariţia”.

Legătura dintre informaţie şi politică este fundamentală şi şi-a pus amprenta asupra vocabularului nostru. Cuvântul *statistică* provine de la termenul german din secolul al XVIII-lea, *Staatwissenschaft*, „ştiinţa statului”, predată de profesorii universitari prinţişorilor Sfântului Imperiu Roman. Într-adevăr, definiţia funcţională a statisticii – „a face ca elemente anterior separate să existe împreună, oferind astfel realitate şi consistenţă unor obiecte mai mari şi mai complexe” – este aproximativ aceeaşi cu scopul politicii, cel dintâi, ca abstracţie numerică, celălalt, ca realitate umană. Iniţial, cuvântul englezesc *classified* [„clasificat” – n.tr.] avea un înţeles

92 Vechile regimuri (în franceză, în original) (n. tr.)

93 Departamente (în franceză, în original) (n. tr.)

principal: sortarea informaţiei pe taxonomii. În secolul al XIX-lea, pe când statul britanic devenea din ce în ce mai puternic, iar imperiul său câştiga în amploare, *classified* a ajuns să primească suplimentar şi înţelesul său actual: informaţie deţinută sub jurisdicţia exclusivă a statului. Chiar şi cuvântul *control* are o origine informaţională, fiind derivat din verbul din latina Evului Mediu *contrarotulare*, care însemna să compari ceva „cu rotulusurile", sulurile de hârtie care serveau drept înscrisuri oficiale.

În *Economie şi societate* (1922), cea mai importantă lucrare sociologică de la începutul secolului al XX-lea, Max Weber anunţa „instrumentul de precizie" al *birocraţiei* ca fiind modul cel mai avansat de organizare a activităţii umane de până atunci. Opinia lui Weber, justificată pe parcursul următoarelor câteva sute de ani, era că un plan unificat de organizare sistematică a informaţiei, bazat pe „precizie, viteză, lipsă de ambiguitate, cunoaşterea dosarelor, continuitate, discreţie (...), unitate" ar constitui cel mai eficient sistem de control politic. Lecţia aceasta nu a fost irosită de Franklin D. Roosevelt atunci când şi-a asumat preşedinţia Statelor Unite, în 1933. Celebrul său program de intervenţie economică a fost asociat cu efortul, mai puţin seducător, dar nu mai puţin vital, de modificare a abordării statisticii de către guvernul federal. Perioada dintre anii 1935 şi 1950 a fost martora apariţiei celor „trei instrumente principale ale statisticii sociale şi economiei moderne": sondajele bazate pe eşantioane reprezentative, sistemul conturilor naţionale şi informatica. Dacă Hamilton, Weber şi Roosevelt ar mai fi trăit încă, toţi ar fi interesaţi de uriaşele cantităţi de date pe care le producem în prezent şi de sistemele tot mai performante pe care le utilizăm pentru procesarea lor.

Făcând un pas înapoi, putem acum să avansăm o ipoteză preliminară care să ne călăuzească pe parcursul cărţii: aceea că modul în care culegem, stocăm, analizăm şi comunicăm informaţia – în esenţă, modul în care o organizăm – este strâns legat de modul în care ne

organizăm viaţa politică. […]

Întorcându-ne din trecut spre viitor, trebuie să ne întrebăm cum vor afecta tehnologiile digitale transformaţionale – tehnologii ale informaţiei şi comunicaţiilor – sistemul *nostru* politic.

Aceasta este întrebarea care stă la baza cărţii de faţă.

Ştim că efectele tehnologiei sunt diferite de la un loc la altul. Apariţia tehnologiei tiparului în China şi Coreea, de exemplu, nu a provocat acelaşi gen de schimbare care a urmat după apariţia tiparului lui Gutenberg în Europa, unde societatea era mai bine pregătită pentru o revoluţie religioasă şi politică. Asemenea diferenţe pot fi explicate, de regulă, prin prisma împrejurărilor economice şi politice. Cine deţine şi controlează o anumită tehnologie, modul în care este aceasta primită de public, dacă posibilele ei întrebuinţări sunt gândite dinainte şi dacă tehnologia este direcţionată spre un anumit scop, toate acestea îi vor influenţa impactul.

Asta înseamnă că ar trebui să nu ne grăbim să tragem concluzia că dezvoltarea unei anumite tehnologii va duce inevitabil sau implacabil la un anumit rezultat social. Gândiţi-vă la internet: pentru că structura reţelei sale era implicit potrivită pentru o organizaţie descentralizată şi neierarhizată, mulţi au prezis, cu toată încrederea, că existenţa on-line va fi destul de diferită de cea din lumea off-line. Însă lucrurile nu au prea evoluat în această direcţie. În principal ca o consecinţă a lumii comerciale şi politice în care şi-a făcut apariţia, internetul a intrat din ce în ce mai mult sub coordonarea şi controlul marilor companii şi al entităţilor politice care ne filtrează şi ne modelează experienţa on-line.

În plus, nu putem să considerăm că tehnologia este echivalentă cu progresul. […]

Pe când Lenin descria comunismul drept „puterea Sovietelor plus electrificarea întregii ţări”, Troţki şi-a dat seama că progresul tehnologic nu constituia un garant al

progresului moral. „Alături de secolul al XX-lea”, scria el, convieţuiesc şi „secolul al X-lea sau al XIII-lea”:

O sută de milioane de oameni folosesc electricitatea, dar continuă să creadă în puterea magică a semnelor prevestitoare şi a exorcismului. (…) Ce rezerve inepuizabile de întunecime spirituală, ignoranţă şi sălbăticie posedă aceştia! (…) Tot ce ar fi trebuit eliminat din organismul naţiunii în (…) cursul dezvoltării nestingherite a societăţii, ni se înfăţişează astăzi, ţâşnindu-i din gât.

Nu putem să luăm de bună niciuna dintre concluzii. Nici nu putem să considerăm că facultăţile morale ni se vor dezvolta automat, odată cu invenţiile noastre. Jocul încă nu s-a terminat.

CUPRINS

Cuvânt-înainte de Ciprian I. Tomoiagă 5
Notă despre titlu 11
Introducere 15
Capitolul 1 – *Puterea* 19
Capitolul 2 – Datele personale informatizate 43
Capitolul 3 – *Justiţia* 75
Capitolul 4 – *Medicina* 110
Capitolul 5 – Automobilele autonome 150
Capitolul 6 – *Infracţionalitatea* 185
Capitolul 7 – *Arta* 225
Încheiere 251
Note 259
Mulţumiri 313
Glosar 315
Index 321

Pentru comenzi şi informaţii, contactaţi:
GRUPUL EDITORIAL CORINT
Departamentul de Vânzări
Str. Mihai Eminescu nr. 54A, sector 1, Bucureşti, cod poştal
010517
Tel. /Fax: 021.319.47.97; 021.319.48.20 Depozit

Calea Plevnei nr. 145, sector 6, Bucureşti, cod poştal 060012 Tel.: 021.310.15.30 E-mail: vânzărieditura@corint.ro Magazin virtual: www. editura corint.ro

ISBN: 978-606-793-598-1

Format: 16/54 x84; Coli tipo: 22
Tiparul executat la:
Hannah Fry

(n. 1984) este matematiciană, conferenţiar universitar doctor şi autoare a unor celebre cârti de popularizare a ştiinţei, în care a tratat aspecte legate de aplicarea matematicii pentru rezolvarea problemelor lumii contemporane, precum şi diversele chestiuni privind utilizarea algoritmilor, a programelor de calculator: *The Mathemotics of Love: Patterns, Proofs and the Search for the Ultimate Equotion* (2015), *The Indisputable Existence of Santa Claus* (2016).

Incit. Pentru a înţelege adevărata ştiinţă din spatele afirmaţiei, citiţi cartea sa, *A Brief History of Everyone Who Ever Lived*: *The Stories în Our Genes* (Weidenfeld & Nicolson, Londra, 2016) (n.a.).

Întreruperea călătorie. Nivelul 5 (*steering wheel opţional*) aparţine automobilelor (precum un taxi robotizat) care nu mai necesită nicio intervenţie umană (n. red.).

dar nu mai este obligat să monitorizeze permanent drumul şi vehiculul. Nivelul acesta seamănă mai mult cu situaţia în care trebuie să convingi un adolescent să spele vasele (n.a.).

** Cope a mai descoperit şi că trebuia să ţină cont de un număr de alte acorduri, pe măsură ce algoritmul funcţiona în continuare. Lungimea unei fraze muzicale şi lungimea unei bucăţi muzicale, de